

Iraq makes even more trouble

The well-substantiated deviousness of Iraq in attempting to conceal its nuclear facilities should remind the rest of us that there needs to be a durable way of regulating engineering exports.

THE government of Iraq appears to be heading for another confrontation with the outside world, this time over the business of nuclear weapons. Last weekend, President George Bush was more outspoken than previously over the difficulty of arranging that inspectors from the United Nations (mostly from the International Atomic Energy Agency) should have access to nuclear facilities in Iraq. Washington has also been musing openly about the possibility that it may be necessary to mount air strikes against Iraq's facilities if access and cooperation continue to be denied. What seems most to have angered the United States is the evidence, gathered by aerial and satellite reconnaissance, that Iraq is continuing to move equipment and materials between plants now believed to harbour nuclear activities of one kind or another. But the underlying insult is that it appears to have come to light that, while Iraq was (before the Gulf War began) still perhaps two years away from the manufacture of some nuclear weapons, it had made more progress towards that end than others had believed.

The temptation to take an even-handed view of what is happening in Iraq should be suppressed. It was a condition of the cease-fire negotiated just a few weeks ago that Iraq's unconventional weapons should be categorized, counted and destroyed before the full rigour of economic sanctions would be abated. The stocks of chemical weapons already found are a great embarrassment to the outside world as well as to Iraq. It is feasible that the large amount of liquid nerve gas so far found could be disposed of by mobile incinerators, but what is to happen to the sarin already loaded into munition, bombs and artillery shells? Endless study in the United States suggests that only robotized disassembly plants can successfully separate nerve gas from high explosives without killing people, but there is only one working plant for this purpose (on Johnson Island in the Pacific). The chances that such a plant could be replicated in Iraq are only small, but such a development would not be welcome elsewhere in the Middle East. That is a nasty and, seemingly, intractable mess.

So is Iraq's nuclear business, and in ways that reflect badly on others than Iraq itself. It seems that Iraq has been producing small amounts of fissile material, not always by the familiar routes of gaseous diffusion, centrifugation and conversion of uranium into plutonium, but also by the electromagnetic separation of uranium isotopes — the technique favoured by E. O. Lawrence during the Manhattan Project. Although, during the Second World War, Lawrence's technique came a little late, there are some good reasons why a

state prepared to wait for nuclear capability might find electromagnetic separation a convenient way of getting it: the technology of the cyclotron is well advertised in the physics textbooks, while old-fashioned power engineering and vacuum techniques are the only technologies required.

But it is unlikely that even such a familiar route to enriched uranium could have been made feasible without the construction of equipment outside Iraq. The same is true of the plants used in Iraq for making nerve gases, at least some of which were supplied by German companies, who obligingly agreed to site them underground. It will be interesting to learn, if UN inspectors catch up with the electromagnetic separation plant, where that equipment was manufactured. The convention that commerce compels secrecy notwithstanding, there is every reason why the United Nations should make public lists of those who have supplied Iraq with dangerous equipment.

Meanwhile, the international community and its technical component should give more thought than it has done to the problems of the supply of potentially dangerous equipment. Parallel with, but independent of, the Nuclear Non-Proliferation Treaty (NPT), there is a solemn agreement between the chief suppliers of engineering equipment that 'they' will not supply non-signatories of the treaty with equipment used for making nuclear explosives. One difficulty is that governments themselves are rarely suppliers, but that domestically registered companies may be. So control must rest on an understanding between governments and private companies not merely that companies will not supply dangerous equipment specified in some list, but that they will also use their best imagination to deny supply of equipment that might be used dangerously, but whose uses have not been specified in advance. That is a lot to ask of recession-driven manufacturers, as shown by the partly British supply of components for Iraq's super-gun that came to light last year.

Is there a way round these difficulties? The ideal would be an agreement requiring all companies to register supplies of major equipment with some technical branch of the United Nations. The understanding would be that information would remain secret, but would be accessible to people whose job would be to scrutinize the pattern of industrial development in quickly developing countries, seeking to ensure that the developments make non-military sense. The experience of the past few months have demonstrated that mechanical engineering concerns are not the only ones to



watch; civil engineers have, for example, done a great deal to strengthen Iraq's arm in the recent war by building bunkers in which to hide Iraq's aircraft, but putting pesticide plants underground is not child's play either. □

What biodiversity?

There is a danger that the concept of biodiversity will become an aspect of political correctness.

WE all so welcome the great variety of living things on the surface of the Earth that the question why diversity is to be encouraged is rarely asked. On the contrary, it is more often assumed that diversity — or biodiversity — is so self-evidently a good thing that the question is both irrelevant and irreverent. But is there not a danger that, like all concepts whose truth is supposed to override dissenting argument, that of biodiversity, like that of motherhood in the United States in the 1950s, may be heading for a tumble? Reports of a meeting organized by the British government's Department of the Environment a few weeks ago (*Nature* 351, 596; 1991) suggest a kind of revivalist meeting from which backsliders might have been expelled (or simply not invited).

So why bother? Because there is no doubt that natural diversity is declining, that there is a powerful movement to halt that process and that it is important to know what the cost of doing so will be. There are several different fields in which the argument for biodiversity may be pitched, and which may be roughly classified as exploitative, ecological, evolutionary, aesthetic and philosophical.

Under the first heading is the argument that less diversity means less opportunity to find helpful medicines in, for example, plant cells; it would be stronger if the dream of making a modern medicine chest from folk remedies had been more amply realized. The ecological arguments are related; do not get rid of the wild mouse, which is a scavenger of scraps of food that might otherwise be breeding-grounds for harmful bacteria, and which itself is food for other nobler creatures, badgers and eagles included. Yet people continue to trap mice apparently blithely neglectful of the ecological imperative that everything hangs together.

Unfortunately, neither class of argument can be held to safeguard two of the creatures on which conservationists have lavished attention over many decades — whales and tigers. Each sits at the top of a food chain independent of that on which people depend (except that Indian tigers often eat domestic bullocks and sometimes even their owners). That is why a stronger case for their conservation is evolutionary — the vivid proof their form provides of the malleability of the mammalian genome and the practical importance of the better understanding of evolution at the genetic level now in prospect. But not even that is clinching; a six-digit vertebrate fossil may be as stimulating of the imaginations of future molecular geneticists as is a visit to a well-run zoo. Comparative physiology deserves a better place than it is given at present, but the present documentation of the natural world should otherwise suffice.

Does the aesthetic argument then take over? Professional people profess themselves uneasy at the idea, protesting that there are no objective yardsticks to hand. But really? It is a matter of common observation that people, given the chance, enjoy the countryside and that the more affluent they are, the more often they make opportunities for themselves to do so. That is neither surprising nor morally wrong, but the biodiversicists would be well-advised to guard the flank they expose to those in the poor countries of the world who have not yet understood why they should keep their squalor intact for the benefit of others.

The philosophical argument is quite different. In the world as it is, human beings are quite successful, but they may not last. So is it not prudent to arrange that some other form of life should succeed to predominance? And since it would be wrong, and speciesist as well, to preordain a successor species, does not inter-species equity require that there should be a level playing-field for the others? Of course. The trouble is that there are no votes in that proposition. □

Washington 'busy'

Last Thursday, Washington D.C. was almost shut down by a computer glitch. But life has continued.

WHAT happens when when a capital city's telephone system breaks down? When a software failure at the telephone company turned Washington D.C. into a giant busy signal one day last week, the *Washington Post* became hysterical and declared, "When the communication system fails, civilization's lifeline seems to snap" as politicians and lawyers, lobbyists and news organizations, and even mothers and daycare centres found themselves unable to make simple telephone phone calls. Ironically, as the telephone system throughout the Washington region went down, telephone company officials were on Capitol Hill lobbying Congress for permission to expand their horizons from telephones to cable television and other communications businesses. Since none of them was on the telephone at the time — cellular telephones are kept out of hearings rooms — the lobbyists did not know that their main business had shut down.

Not to be outdone, a software glitch at Pacific Bell, the California telephone company, put Los Angeles on hold for three hours on the same day, apparently by coincidence. The basic problem seems to be that the telephone industry is offering more services to more people than technology can sustain. In Washington, a new routing technology called Signalling System 7 allows more calls to be completed more quickly than ever before, while allowing customers to buy add-on services such as "Caller-ID" that displays the caller's number in advance.

Especially in Washington, there has been much tension. Psychologists (routinely consulted by the press in emergencies) report that when telephones are out of order, "people tend to get aggressive very fast". It is, they say, a problem because nervous systems are trained for instant fulfillment from telephones. □

Slow progress in Geneva on warming treaty

- US still resists emissions targets
- 'Pledge and review' could break impasse

Washington

THE international negotiating team drafting a treaty to limit global warming is trying desperately to find language that the United States — alone among the developed nations in refusing to declare a target date by which its emissions of carbon dioxide will be stabilized — can support. But progress is slow, and the second two-week meeting of the United Nations climate convention negotiating committee ended in Geneva last week without agreement on how to limit the build-up of greenhouse gases in the atmosphere.

The aim of the negotiations, which continue in Nairobi in September, is to produce a climate convention for signature at the United Nations Conference on Environment and Development, scheduled for June 1992 in Brazil.

Most observers agree that an international agreement to reduce world greenhouse gas emissions below their present level (said by last year's Intergovernmental Panel on Cli-

mate Change to be necessary if the global mean temperature is to be stabilized) is impossible before 1992. But many countries would like to see the initial convention include an intermediate commitment to limit greenhouse gas emissions, rather than being a simple expression of concern about global warming. Many European Communities governments believe a feasible target would be for the industrialized countries to agree to stabilize their carbon dioxide emissions at present levels by 2000.

This is where the United States' position is a stumbling block, and an important one because the United States emits a greater volume of greenhouse gases than any other country. The US administration will not countenance a carbon dioxide emissions target, stating that total US emissions of greenhouse gases in 2000 will be no greater than in 1987 (see *Nature* 348, 4, 1990). But this includes controls on the use of chlorofluorocarbons made under the Montreal Protocol — cuts most nations say cannot be 'double counted' under a new climate agreement.

To break the impasse, Japanese, British and French negotiators in Geneva proposed a system of 'pledge and review' of greenhouse gas emissions. Under this proposal, countries signing the convention would produce a statement detailing a strategy to control emissions. Progress in each country will be reviewed periodically by a team of international experts.

The wording was sufficiently vague to avoid alienating the United States at this stage, allowing countries to produce their own plans to control emissions, which need not include specific commitments on carbon dioxide. Environmentalist groups in Geneva were quick to attack the pledge-and-review proposal, labelling it "hedge and retreat". And on the last day of the meeting, the Dutch delegation, speaking for the European Communities, said that pledge-and-review is only acceptable if the pledges take the form of binding commitments — something unlikely to win US support.

With no other firm proposals on the table, pledge and review is not yet dead in the water. But if the climate convention is to have any real teeth, it may be necessary to set an overall emissions target for the pledge and review process to meet, and for pressure to be put on those countries that fail to pull their weight. How far this pressure could be applied without the United States (and probably a number of developing countries) rejecting the convention on grounds of an infringement of national sovereignty remains to be seen.

Peter Aldhous

HUMAN GENOME PROJECT

HUGO flirting with Johns Hopkins

Washington

THE Human Genome Organization (HUGO) is negotiating to affiliate its American office in Bethesda, Maryland with Johns Hopkins University, less than one hour away in Baltimore. Such a move would allow the office to receive grants from federal agencies, ending its reliance on charitable funding.

HUGO's Bethesda office is now funded through a four-year, \$1-million grant from the Howard Hughes Medical Institute. But if HUGO is to fulfil its intended role of coordinating the international effort to map and sequence the human genome, some support from governments is essential. The problem in the United States is that institutions applying for federal grants must conform to set standards for accounting, social security arrangements for their staff and the like. HUGO, with limited funding, has been unable to do this. "We're in a catch-22 situation," says Diane Hinton, who runs the Bethesda office single-handedly.

Affiliation to Johns Hopkins may be the answer. HUGO could then apply through the university for federal grants to support its Bethesda office. Hopkins officials have said they can see no legal obstacles to the affiliation, and Hinton and HUGO director James Wyngaarden visited Baltimore last week to begin the negotiations proper.

Any large university could serve as an umbrella organization for the HUGO office, but Johns Hopkins is particularly appropriate. Apart from its proximity to Bethesda, Hopkins also houses the Genome Database, the main repository for human gene mapping data.

Since HUGO was conceived in 1988 as an organization of genome researchers to provide 'bottom-up' coordination between the various national genome initiatives, one goal has been to win government funding. But HUGO's regional offices have so far had to depend on charitable support.

The European office in London has got off the ground with help from the Wellcome Trust and the Imperial Cancer Research Fund, and there are hopes that European governments will also provide money. But the Pacific office in Osaka has been paralysed by a lack of funds. No government agency has shown any interest in providing money for HUGO and the Japanese office is finding it difficult to raise money from private sources because of government red tape that inhibits the establishment of tax-free foundations.

Ironically, HUGO's newest venture — a local office in Moscow that opened this week — is the first to get official government backing, with the Soviet authorities pledging financial support.

Peter Aldhous

Power lobby pressure

A COALITION of US coal companies and electricity utilities is considering a national advertising campaign later this year to reduce public support for energy policies designed to slow global warming. The campaign, organized by a new organization called the Information Council for the Environment (ICE), has already been tested in a number of small US cities. "How much are you willing to pay to solve a problem that may not exist?" asks the headline on a sample full-page newspaper advertisement.

ICE is advised by a number of the noted greenhouse sceptics among the scientific community, including Patrick Michaels, from the University of Virginia, and Robert Balling, from Arizona State University. Its board is now mulling over the response to the trial campaign before deciding whether to go national.

The campaign has angered environmentalists, who accuse the power lobby of disinformation, ignoring the scientific consensus summarized in last year's report from the Intergovernmental Panel on Climate Change. In particular, green activists dispute one advertisement that claims Minneapolis is getting colder, pointing to a 1989 University of Minnesota study showing a significant rise in mean temperature in Eastern Minnesota since the 1850's.

If the campaign had run as originally suggested by ICE's advertising agents, the environmentalists' ire may have been even greater. One suggested headline, thrown out by ICE's board stated: "Some say the Earth is getting warmer...some also said the Earth was flat."

P.A.

Protecting US biotech firms

Washington

Two bills have been introduced into Congress that are intended to reform US patent and trade law and protect the US biotechnology industry against unfair foreign competition. The proposed legislation would bring US patent law in line with that of Japanese and Western European nations, and afford US inventors the same proprietary protection enjoyed by their foreign competitors.

The two nearly identical bills, introduced by Dennis DeConcini (Democrat, Arizona) in the Senate and Rick Boucher (Democrat, Virginia) in the House of Representatives, would remedy the confusion caused by a 1985 court decision, which severely restricts the ability of the US Patent and Trademark Office to award US biotechnology companies with process patent protection. If enacted, the legislation would overturn the earlier court ruling and provide the patent office with the authority to issue process patents as long as the starting materials are "novel".

The proposed legislation would "provide a simple solution to a complex area of law", DeConcini says.

Current US patent law works against biotechnology companies because it denies a patent on anything that is not "novel" or that is considered "obvious" in light of existing knowledge and techniques. Given these requirements for patentability, most biotechnology products fail the test.

The problem arises because biotechnology, unlike the traditional pharmaceutical industry, is not producing anything new.

Biotechnology involves the use of a known process — recombinant DNA technology — to produce a genetically engineered version of a naturally occurring protein or enzyme. Although biotechnology is often the only means of producing practical quantities of a given material, product patents on a recombinant version are routinely denied by the patent office on the grounds of lack of novelty.

Compounding the problem is a 1985 federal circuit court decision, *In re Durden*, which has made it increasingly difficult for biotechnology companies to obtain process patent protection for biotechnology inventions. In that case, the court held that process patents could not be issued for processes that use novel starting materials — which are patentable in their own right — in combination with a known chemical process. Although the *Durden* case was itself not directly related to biotechnology, the patent office has taken the case into account when making its rulings on biotechnology process patents.

Supporters of the proposed legislation argue that many of the patent problems for biotechnology stem from what they see as an erroneous and inconsistent application by the patent office of the *Durden* decision.

Without the benefits of process patent protection, domestic companies are unable to prevent foreign companies from exploiting a legal loophole that exists in US patent and trade laws. According to those laws, if a product made by a patented process is manufactured offshore and imported, it constitutes infringement, but if there is no process patent involved, there are no restrictions on the importation of the products — even if they were made with a starting technology, such as a gene or host cell, that is patented in the United States.

As long as biotech companies have a difficult time obtaining process patents, they will be vulnerable to unfair trade practices, say the bill's sponsors.

Proponents of the bill argue that a special case should be made for the biotechnology industry. The scale of investment in both time and money involved in bringing a new biopharmaceutical product to market, which typically takes 10 years and costs between \$100 and \$200 million, "stands in stark contrast to the ease at which the product can be copied", says Henri A. Termier, president and chief executive officer of Genzyme Corporation.

The bill has bipartisan backing within Congress and the support of the Administration. It has also been endorsed by the major trade associations representing the biotechnology industry and has garnered wide support among universities. Although *Durden* rejections can sometimes be overcome, the bill's sponsors believe that smaller biotechnology companies and universities often forfeit process patent protection because they lack the financial resources to challenge the patent examiner's decision.

Opponents of the bill caution against congressional intervention at this time, particularly as the corrective legislation would apply to all technologies and not just biotechnology.

Given that biotechnology is still a young industry, the law in this area should be allowed "to mature through court decisions based on thoroughly presented factual situations without premature legislative action", says William F. Marsh, speaking on behalf of the Intellectual Property Owners Inc., a non-profit association representing patent owners among the technology-based industries.

Marsh argues that as the bill would require the patent office to issue process claims without examination for novelty and non-obviousness, it would have a "destabilizing influence" on the US patent system. Far from clarifying the problems that have arisen from following the precedent of the *Durden* case, enactment of the legislation would lead to "great uncertainty as to the validity and scope of the process claims after the patent is issued", Marsh says.

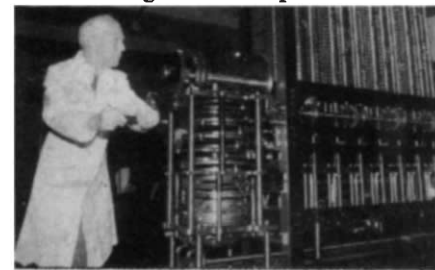
Diane Gershon

Cranky computer

London

With a turn of a handle last week, a three-tonne mechanical computer designed almost a century and half ago correctly cranked through its first public calculation: $0+0=0$. Known as the Difference Engine No. 2, the computer has been constructed by the London Science Museum to answer the question of whether Charles Babbage could have gone ahead and built the world's first computer after he designed it in 1849. As far as adding zeros go, Babbage seems vindicated.

But more challenging calculations still trip up the gear-driven behemoth. Two times two sometimes comes out one hundred trillion and four, a problem museum computer curator Doren Swade attributes to a faulty mechanism that carries "ones" to the next higher decimal place after addition overflows.



Of the 210 such mechanisms in the steel and bronze machine, "we've got a few rogue carries," he said.

In private testing, the difference engine correctly calculates a table of numbers raised to the power of seven most of the time. But that is still not good enough; the machine is destined to be the top attraction of a history of computing exhibition at the science museum, and to settle the Babbage dispute. Engineers on the £300,000 project are now trying to get the machine running well enough for reliable calculations. But the machine was not designed for easy adjustment and Swade says it is possible that it may never reach full operation.

C.A.

ZOO EXCHANGE

Kuwaiti home-from-home

London

If the London zoo, facing a £2 million shortfall, is forced to make good its threat to close down this autumn, it will find at last one eager taker for its animals. Moussa al Khasti, director of the Kuwait zoo, which was ransacked during the Iraqi occupation, last week offered to take the London collection to replace his own animals, some of which were eaten by Iraqi troops. Only the Kuwait zoo's elands and buffaloes appear to have actually been eaten by the Iraqis, al Khasti says, but hundreds of birds, dozens of mammals and all 40 reptiles save three tortoises were either released or killed during the conflict. C.A.

INDIRECT COSTS

OMB announces cap

Washington

THE White House Office and Management and Budget (OMB) has revealed the details of its proposed cap on the rates that US universities can charge to federal grant-making agencies to cover the overhead, or "indirect", costs of research. As expected, the new rules would limit the portion of overhead charged to cover universities' administrative costs to 26 per cent of the value of research grants — about the average current rate.

The 26 per cent cap comes into force for any overhead contracts negotiated or amended after 30 September, and is expected to cost US universities at least \$75 million over the coming year.

Robert Rosenzweig, executive director of the Association of American Universities, is concerned about this compressed timescale. "There's a case to be made for phasing the cap in for institutions that are going to be hit hard in the short term," he says. For universities that now receive high rates of administrative overhead, those rates may have been won in the give-and-take of negotiations at the expense of a lower reimbursement for the costs of buildings and equipment. Capping

the administrative overhead without modifying the other parts of the contracts could be unfair to the schools.

Universities have until the end of July to comment on the proposed rules. Given the current scandal over indirect costs, few observers expect major changes in response to university protest.

Michael Hall, a staff member at the Senate appropriations subcommittee responsible for the National Institutes of Health's budget, says that any controls coming from Congress would probably be even stricter.

Rosenzweig concedes that, for the time being, the universities will have to live with a 26 per cent cap on administrative costs.

But he wants federal agencies to agree quickly on a new system of accounting for and reimbursing research overheads, so that the cap does not become a permanent fixture. Rosenzweig's association had been negotiating with the OMB to reform the system, before congressional outrage earlier this year at the inclusion of a yacht and other luxuries in university research accounts at Stanford University forced more drastic action by the OMB.

State universities are particularly concerned about a second provision in the OMB's proposals demanding that money reimbursed for depreciation of buildings and equipment is actually used to cover future building and equipment costs. University officials liken this to an employee submitting an expenses form for a business lunch, and then being forced to spend the reimbursement on another lunch. The lack of budgetary flexibility, says Bill Brophy, a finance officer responsible for indirect costs at the University of California, San Diego, will make it more difficult for state universities to return money owed to their state legislatures.

Peter Aldhous

■ In a separate development last week, the Office of Naval Research announced that it has disciplined six of its civilian employees, responsible for negotiating the indirect cost rate on research grants to Stanford University.

The action against the six civilians ranges from assignment to another job and a cut in pay, to a simple letter of admonishment.

The Navy's move is the first action taken against federal officials who failed to prevent the charging of luxury items to university research accounts. Stanford alone has now returned over \$1 million incorrectly charged to research overhead — a minute fraction of the total sum thought to be charged incorrectly by Stanford to the US taxpayer over the 1980s.

TECHNOLOGICAL RESEARCH

US losing its lead

Washington

THE United States' world lead in technological research is diminishing, according to an analysis of the Institute for Scientific Information's Science Indicators database, published in the latest issue of *Science Watch*. The average citation impact of papers with US authors in most fields rose through the 1980s. But in engineering, technology and applied sciences, those research areas most closely linked to industrial competitiveness, the citation of US papers has fallen by 6.9 per cent, relative to the world average. The sharpest decline was in metallurgy — 20.3 per cent, nuclear engineering — 18.1 per cent, and instrumentation — 13.6 per cent.

Somewhat surprisingly, the declining impact of US technological research papers is not due to a corresponding rise in citations of papers from Japan, whose technologists are no more likely to be cited now than in the early 1980s. Instead, it seems that it is France and Britain that are the growing forces in engineering, technology and applied science. But US technologists should not panic.

Despite the decline in citation impact over the 1980s, US technological research papers are still cited over 60 per cent more often than the average for those from the G7 group of nations.

Peter Aldhous

MATERIALS SCIENCE

Superconductivity hot again

Washington

ADVANCES in high-temperature superconductivity have been so rapid over the past year that the use of superconductors in electric power systems, large motors and other energy applications is likely to come much sooner than most expected.

For that reason, the United States should form a \$250-million, five-year programme to develop electric energy systems that use high-temperature superconductors. So says a report issued late last month by a group of superconductivity specialists from industry, academia and national laboratories.

"The progress has been astonishing," says Alex Malozemoff, vice president of research and development at American Superconductor Corporation in Watertown, Massachusetts, and a member of the group that wrote the report. "In the last year it has been confirmed that one can actually build things with these materials."

After the initial euphoria accompanying the discovery of high-temperature superconductors died down, researchers found several problems with the materials that threatened to slow the development of 'bulk' applications — such as wires, magnets and motors — to a crawl. The superconductors were brittle and inflexible, which made them poor candidates for wires that would be wound in coils, and they exhibited a variety of interrelated problems that limited their critical current density — the amount of current they could carry and still remain super-

conducting.

But the progress in increasing the critical current density has been faster than anyone expected, and American Superconductor and other labs have shown they can make flexible superconducting wire. Suddenly, people are again excited over the potential applications of high-temperature superconductors to such devices as magnets, motors, generators and electrical transmission wires.

That creates an opportunity for US industry, the report says.

Japanese and European researchers are developing superconducting power systems that use the traditional low-temperature superconductors. "With an aggressive program, the United States could leapfrog its competition by demonstrating the components of a high-temperature system within the next five years," the report predicts. One of the most valuable applications would be superconducting motors, it says.

The five-year programme envisioned in the report would be led by the federal government but with significant private-sector investment. For each class of applications, the report recommends establishing "vertically integrated, industry-led teams with representation from manufacturing, universities, national laboratories" and end users of high-temperature superconductor such as electric utilities. The development would take place at "a pre-competitive level but with a long-term, market-driven focus."

Robert Pool

The greening of Collor

São Paulo

For the first time, Brazil has agreed to accept "debt-for-nature swaps", in which up to \$100 million in outstanding loans will be retired each year in exchange for government spending on the environment.

The news from Brazilian president Fernando Collor de Mello came last week along with the announcement of several other environmental initiatives. Collor had just returned from the United States, where he was publicly criticized by US congressmen and environmentalists for not doing enough to curb devastation of the Amazon rain forest. These moves may help turn his image a shade greener.

Debt-for-nature swaps were anathema to the previous government of José Sarney, but with Brazil's external debt at \$123,000 million, the new government has reconsidered.

A key figure in the change of policy has been the new economy minister, Marcilio Marques Moreira, former ambassador to the United States, who has been more sympathetic to the environmental movement than many other Brazilian officials. In January 1989, for example, he participated in a small ceremony in front of the US Capitol to honor the environmental activist Chico Mendes, who had been murdered the month before.

At the time, his actions seemed out of line with the tough stances taken by other offi-

cials, such as former foreign minister Abreu Sodré, who once said, "Brazil won't be the world's ecological preserve."

In the swaps announced by Collor, environmental groups or other governments will buy outstanding Brazilian debt on the secondary market, where it trades for little more than 30 per cent of its face value. The Brazilian government will then "pay off" the debt with local currency, putting about 75 per cent of the face value into accounts that will be used for various environmental purposes.

The money will be used to survey park boundaries, train guards and demarcate parks and Indian lands; some may go to research institutes. Belém's Museu Emilio Goeldi, for example, is barely functioning and part of its exhibition was closed to the public for lack of funds.

Along with the announcement of the swaps, Collor signed a decree forbidding tax subsidies for farm and cattle ranching projects that destroy the environment. The subsidies had been stopped in the beginning of his term but were re-initiated in April after pressure from supporters of the government.

Collor also fired Cantidio Guerreiro Guimarães, the head of Funai, the Indian protection agency.

Collor said the reason for the firing was that the agency head had failed to demarcate the lands of the Yanomami Indians, the largest tribe in the country, in the state of Roraima, near the border with Venezuela.

Cantidio Guerreiro says he was a scapegoat, fired to placate foreign critics and relieve international pressures. But he was considered to be a passive leader by those agency workers who live among the Indians and try to protect them from miners and farmers. The agency's number two man, Edivio Batistelli, who will succeed Cantidio Guerreiro, has a better record of dealing with the Indians and is considered to be a better negotiator.

Collor's measures strengthen the position of his secretary for the environment, José Lutzenberger, who was chosen from the ranks of the country's environmentalists. The secretary had been criticized by some as a mere ornament in Collor's cabinet — a figurehead to appease foreigners. Now Lutzenberger will have his chance to prepare the country for next year's international meeting on biodiversity and the environment, to be held in Rio de Janeiro in June. It will be a chance for Brazil to cast off its image of international environmental villain — a reputation the country acquired with the extended burnings in the rain forest in 1988 and furthered by the killing of Chico Mendes.

This year's rainy season has just finished, and the dry season — suitable for burnings of forest areas — is about to begin. The next few months will tell much about Collor's commitment to the environment.

Ricardo Bonalume

More RU-486 now available for research

Washington

ARE anti-abortion activists keeping the drug RU-486 out of the hands of researchers who want to test it for non-abortion uses? Absolutely, said organizers of a new organization created last month to oppose the current import ban on RU-486 for abortion use.

"The majority [of citizens] want RU-486 brought into the country for basic research. It's outrageous that you can't even test it," said Carol Rowen, a spokeswoman for the Los Angeles Coalition for RU-486, a collection of two dozen Californian health, abortion and women's groups who announced their joint campaign at a press conference.

But the coalition may have leapt before it looked. Although RU-486 availability may have once been a problem, it has been freely available to researchers for months, scientists say.

In support of their campaign, coalition organizers cite a congressional hearing by US Representative Ron Wyden (Democrat, Oregon) last November as proof of the RU-486 availability problem. At that hearing, researchers testified that they were finding it difficult, if not impossible, to obtain supplies of the drug from its French manufacturer, Roussel-Uclaf, because anti-abortion protesters had threatened the company with a boycott if it imported the drug into the United States (see *Nature* 348, 382; 29 November 1990).

Since then, however, the outlook for RU-486 research has brightened considerably, something the coalition's organizers appear to have neglected to check out before launching their nationwide campaign.

"We don't have any problem anymore. We can get all we need," says George Chrousos, a National Institutes of Health researcher studying RU-486's use for Cushing's syndrome and one of the witnesses at the November hearing who testified that they had been unable to secure a supply of the drug. Other scientists who are studying RU-486 for breast cancer and other diseases agree: whatever the extent of the problem last year, neither Roussel-Uclaf nor the US government appears to be standing in the way of legitimate research now.

Coalition organizers said they would look more closely into the issue. If they are unable to find a scientist with a complaint about RU-486 access, they say they will narrow their focus to the current import ban on the drug as an abortifacient. On that front they may find support easier to find. This spring, New Hampshire offered itself as a haven for RU-486 clinical testing for abortion, and a bill now pending in California would do the same.

Christopher Anderson

LIBRARIES

Museum still looking for stolen books

São Paulo

THREE rare books stolen from the Museum of Zoology of the University of São Paulo have been recovered, but 26 others are still missing. The books, some of which are worth as much as \$200,000, all deal with birds and avian taxonomy.

The theft occurred on 8 March, but three of the books were found a month later at the auction house Christie's in New York, but the Brazilian press only recently learned of the case. The three books were found because a Christie's expert, Stephen Massey, thought it strange that they carried stamps of the Museu Paulista, the institution to which the Museum of Zoology was affiliated up to 1939.

The US Federal Bureau of Investigation and Interpol, the international police network, are looking for the other 26 books.

Brazilian police suspect that there is a gang specializing in stealing rare books from the badly protected libraries in this country. The books may be identified by stamps that say "Museu Paulista", "Museu de Zoologia" or "Universidade de São Paulo".

Ricardo Bonalume

NUCLEAR POWER

Japan debates plutonium

Tokyo

WHAT should Japan do with a vast pile of weapons-grade plutonium that the Japanese nuclear power industry will have on its hands by early next century? After two years of debate, a subcommittee of the Atomic Energy Commission is expected to recommend soon that most of the plutonium — enough to make thousands of nuclear weapons — should be used in conventional nuclear power plants until commercial breeder reactors, which are designed to use plutonium, come into operation sometime in the distant future. Critics say, however, that this is a very uneconomic solution and that instead Japan should completely revise its nuclear policy to eliminate production and accumulation of the highly toxic and dangerous material.

The plutonium — derived from reprocessing of spent Japanese nuclear fuel — will be brought back into Japan from reprocessing plants in France and the United Kingdom beginning next year. The imports themselves have been a matter of international controversy because of possible accidents or hijack during transit, and Japan has had to build a special armed vessel for the purpose.

In addition, a new commercial reprocessing plant that is being constructed in the north of Japan will also supply the industry with plutonium beginning in 1998.

Severe delays in development of commercial breeder reactors have left Japan with the problem of putting the plutonium to practical use. When the Japanese nuclear power industry signed contracts many years ago with the United Kingdom and France to reprocess nuclear fuel, Japan was planning to use the plutonium in commercial breeder reactors by 2010. But technical difficulties in development of breeder reactors and the plentiful supply of cheap uranium for con-

ventional reactors have pushed back the starting date for commercial breeder reactors to sometime between 2020 and 2030. And there are doubts that even this schedule will be met.

Atomic Energy Commission officials estimate that Japan's inventory of plutonium will balloon to 84 tons by 2010, compared with only a few tons at present, according to a report from the Kyodo wire service last week. The commission's nuclear fuel subcommittee is expected to recommend that most of the plutonium be used in twelve conventional light water reactors by mixing it with ordinary uranium fuel.

Kimikazu Iwase, deputy director of the nuclear fuel division of the Science and Technology Agency, says that the Kyodo figures are not the official ones that will be released in the next few weeks. The subcommittee's final figures may be "slightly different", he says. But the conclusion that the plutonium should be used in conventional reactors is in line with the commission's 1987 long-term policy, he adds.

Paul Leventhal, president of the Washington-based Nuclear Control Institute, an organization that monitors proliferation of nuclear weapons, questions Japan's whole policy. "Using plutonium to fuel ordinary light-water reactors is inefficient and dangerous," he says. "The introduction of plutonium-uranium fuel is just a stop-gap measure to avoid an enormous plutonium surplus. The best way for Japan to avoid this worrisome surplus is not to create it in the first place." Leventhal recommends instead that Japan's nuclear power industry renegotiate its reprocessing contracts with France and the United Kingdom to take back low-enriched uranium instead of plutonium.

Iwase admits that plutonium is much more expensive than uranium and thus at present

is not a very economic fuel for conventional nuclear reactors. But he argues that in the long term as Japan uses more and more plutonium, the price will drop.

Price, however, is not Japan's main concern. Government officials and the electric power companies are determined to establish a complete nuclear fuel cycle in Japan, including the production of plutonium from spent nuclear fuel, so that the nation will be less dependent on fossil-fuel imports. The government and industry have already invested thousands of millions of dollars in this policy, and it seems they are determined to forge ahead no matter what the cost.

David Swinbanks

ARMSCONTROL

South Africa to sign

London

AFTER years of suspicion that South Africa was developing nuclear weapons, its government last week ended the speculation by both admitting that it had the capability to produce such weapons and at the same time promising never to use them. In agreeing to sign the 1970 Nuclear Non-Proliferation treaty, President de Klerk said that South Africa would open all its nuclear facilities to outside inspectors and move towards establishing the country as a nuclear-free zone. The treaty essentially promises that the signatories will neither acquire nuclear weapons if they do not have them, nor help non-nuclear nations gain nuclear capability. In the aftermath of the Cold War and the reduction in superpower tensions, several previously recalcitrant countries — including France last month — have agreed to sign the treaty. Of the nations that had nuclear capability when the treaty was established, only China is yet to sign.

C.A.

NUCLEAR POWER

Bulgaria's dilemma

Munich

IN an unusually strong statement, the International Atomic Energy Agency (IAEA) has urged the Bulgarian government to shut down its nuclear power plants in order to avoid a catastrophic accident.

The Vienna-based United Nations agency warned the Bulgarian government last month that its plants were not safe and that they should be shut down immediately, an agency spokesman said. Bulgaria responded by shutting down two of the four plants located at Kosloduj on the Danube River, but it continues to operate two other plants because it said it could not afford to give up the power they provide. All four are 440-MW plants of the type WWR-440 built by the Soviet Union.

Concerned Western officials, led by German environment minister Klaus Töpfer, have agreed to meet at IAEA headquarters in Vienna next week to discuss providing alternative sources of energy to Bulgaria.

S.D.

Monju fast-breeder reactor completed

FAST-BREEDER reactors may have an uncertain future, but officials of Japan's Science and Technology Agency were all smiles a few weeks ago as they celebrated completion of construction of Monju, Japan's first prototype fast-breeder reactor.

Located on the Japan Sea coast, Monju represents one of the biggest investments ever made by the Japanese government in a research and development project. Six hundred thousand million yen (\$4,300 million) has been sunk in Monju. Construction of the plant, which took nearly six years, required a workforce of more than 2,600 and involved four major equipment manufacturers (Toshiba, Hitachi, Fuji Electric and Mitsubishi Heavy Industries) and nine construction companies.

Monju has a loop-type primary and secondary cooling system filled with

liquid sodium. The sodium provides very efficient heat transfer to the steam generator, but it is an extremely dangerous material which catches fire spontaneously in air or in the presence of water, and much of the cost of construction has gone into making the sodium cooling system leak-proof. The reactor has also to withstand major earthquakes.

The prototype fast-breeder reactor, which will have an electrical output of 280 MW, is scheduled to reach criticality in October next year, and Japan plans to begin imports of plutonium from Europe at that time to fuel the reactor. The next step will be for the electric power industry to build an even bigger demonstration fast-breeder reactor. But details of that project are not yet known, and Japan is not expected to have commercial breeder reactors in operation until at least 2020-30.

D.S.

Uranium mining's legacy

Munich

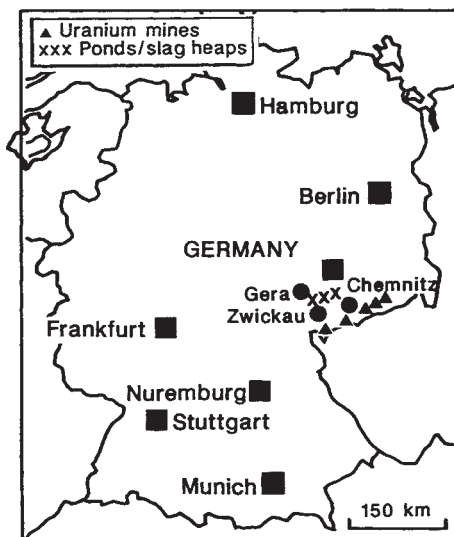
THE costs to human health and the environment of 45 years of uranium mining in eastern Germany are becoming clear as a unified Germany contemplates cleaning up the radioactive mess left behind.

The recently terminated mining activities, which provided the Soviet Union with most of the uranium it needed for its early atomic bombs, have left behind a devastated wasteland close to heavily populated areas.

This otherwise dark cloud does, however, offer at least one small silver lining. It may provide researchers with one of the best opportunities since Hiroshima to learn about the effects of radiation on humans.

Making the uranium mining areas once again safe for human habitation will probably be the costliest clean-up in German history. The expense has been estimated at between DM 5,000 million and DM 15,000 million (\$2,800 million \$8,300 million), although officials of the Environment Ministry say privately that it may ultimately cost much more than that.

The human costs have been equally severe. According to official figures, at least 5,200 miners died from "Schneeberg lung disease", a form of lung carcinoma named after one of the uranium mining towns, which are found in the Erzgebirge mountains in Saxony and Thuringia along the southern rim of what used to be East Germany.



But no one knows if this figure is accurate, and some researchers think many more people have died. Munich physicist Eckhard Krüger, who visited the region recently, estimated last month that there could have been as many as 15,000 radiation deaths among the miners and the local population, but he admits that this is just a guess.

The German Environment Ministry is now looking into the matter. Recently, the ministry launched a preliminary investigation to determine the health effects of radiation on mine workers, who are thought to

have been endangered most. The study is expected to be completed by the end of August.

Making an accurate assessment of the health effects on mine workers — for epidemiologists the most interesting group — will turn on the answers to a few critical questions. How accurately can officials now estimate the radiation dose received by the individual miners? To what extent can medical records from the region be trusted? And to what extent can the natural background radiation in the region be determined? For the crucial period between 1946 and 1955, when miners were provided with almost no protection from radiation by the Soviet owners of the mines, there are virtually no individual dose data, says Hans-Henning Landfermann, a radiation protection official of the Environment Ministry in Bonn. Nevertheless, researchers hope that, by using data on overall radiation levels in the areas where the miners were working, at least the gross effects can be measured.

The medical records from the region are also the object of some controversy. Doctors working under the Communist government were expected to file reports that were not alarming, says Edmund Lengfelder, a radiation biologist at the University of Munich, and therefore the reports did not accurately reflect the true health of the population.

Landfermann disagrees about the extent to which records were falsified. At least for the miners, he says, company doctors at Wismut AG, the jointly owned Soviet-East German company that operated the mines and was taken over last year by the German Economics Ministry, were willing to recognize lung cancer as a work-related illness.

Determining the natural background radiation in the region presents still other problems. Because the region where the mines are found was also mined heavily for silver, nickel and other metals, uranium ores have been dug up and discarded at least since the fifteenth century, says Dietmar Weiss of the Environment Ministry's regional office in Berlin. So the background levels of radiation could have been high even before 1945.

Other influences may have as much to do with the health problems of the people in the region as radiation doses. Ministry officials found that tonnes of arsenic and heavy metals had been unearthed because of both silver and uranium mining. Furthermore, in the early years, when as many as 500,000 miners lived and worked in the region, they were treated to fringe benefits such as extra cigarettes that may well have contributed to a rapid decline in their health.

The government recently admitted in effect that the effects of the mining were more severe than previously thought when it began to compile a new registry of waste dumps and other sites affected by mining operations.

Steven Dickman

Call for brain-maps

AN Institute of Medicine report, published in Washington last week, calls for \$10 million a year to be set aside over five years to begin developing three-dimensional computerized maps of the human brain, as well as of the brains of key laboratory animal species. The initial request is small, requiring an addition of less than one per cent to the annual US neurosciences budget. But the entire project, conceived as a 20-year initiative, is the most ambitious biological informatics programme proposed to date. The construction of databases to hold human gene mapping and sequence data, now underway under the auspices of the Human Genome Project, is a minor problem in comparison.

As envisaged in the report, brain researchers will be able to explore three-dimensional computer maps of the brain, focusing on a particular area and calling up a wide range of information from linked databases about the region in question — including function, pharmacology, and connections to other brain areas. P.A.

Californian quake

THE largest magnitude earthquake to hit California since the Loma Prieta event of 1989 left two dead last Friday (28 June). The 5.8 magnitude quake, centred 11 km beneath the San Gabriel mountains north of Pasadena, is the first with a magnitude greater than 4.5 to be recorded on that region of the Sierra Madre fault since seismologists began monitoring in the 1920s. Lucy Jones, from the US Geological Survey at Pasadena, says there had been some debate as to whether the fault was dead. "Obviously it's not," Jones adds that the latest quake is the fourth around the San Gabriel mountains since 1987. The combined effect of these events is to rotate the San Gabriel valley in a clockwise direction, she says. P.A.

Eastern European aid

THE UK government has added £5 million to its aid for Eastern Europe to start a new exchange programme for environmental researchers. As part of the existing £30 million "Know How" economic aid programme, the three-year exchange effort will bring Eastern European researchers to Britain to learn environmental restoration techniques, and send UK researchers to Eastern Europe to work alongside scientists there. In announcing the new programme, UK environment secretary Michael Heseltine said he had learned — "not always willingly" — the importance of non-governmental organizations like Greenpeace in raising public consciousness towards environmental problems. One of the Know How projects will bring leaders of Czech and Slovak non-governmental organizations to Britain next month to work with their counterparts. C.A.

The animal rights question

SIR — It was good to read Barbara Culliton's clear and critical appraisal of the problems faced by the medical research community as a result of increasingly hostile campaigns by animal rights groups (*Nature* 351, 517; 1991). Although there has been organized protest against animal research for over 100 years, during the past two decades we have had to deal with the far more aggressive animal rights movement. Since the mid-1970s, there have been enormous improvements in the welfare of laboratory animals and controls on their use, but the same period has seen protests against their use escalate from fringe status to a movement embracing both outright terrorism and multimillion pound campaigning organizations.

Nobody would deny that we have a responsibility to preserve the well-being of animals in our charge; the welfare of laboratory animals is a primary concern for the researchers and technicians who work with them. This stands in clear contrast to the idea of 'animal rights', which argues that other sentient animals have similar rights to humans, making it morally wrong for humans to use them in any way: in agriculture, zoos, sport, research or even as pets.

Without animal studies, fundamental physiology, probably the earliest experimental medical science, could not have made significant progress. Immunology and pharmacology could not have become established without experiments on mice, rats and rabbits. Animal studies continue to play a crucial part in genetics, toxicology, pathology, endocrinology, neurobiology and developmental biology and surgery. The huge advance in our ability to diagnose and treat disease over the past 50 years has, to a great extent, depended on animal research.

Yet despite this, new legal controls on how and why animals can be used in research were introduced in Britain five years ago, and despite the fact that they are regarded internationally as the most comprehensive and exacting anywhere, the animal rights movement continues to push for tougher legislation. The truth is that, in Britain, the controls are so tight that administrative delays are already threatening to hold back progress in some areas of research.

For many years, the general principle of "keep your head down and it will go away" dominated the thinking on how to deal with this problem. But now there is a growing conviction in the scientific and medical community that the problem has become too serious to ignore. The general public is being actively misled by the animal rights movements into believing that animal research is both cruel and pointless. Unless the medical and scientific community makes greater efforts to explain to the press, the public and politicians why animals are used in research and the benefits which have come from that use, we can look forward only to greater pub-

lic antagonism, more hostile reporting and, eventually, more restrictive legislation.

Culliton is concerned chiefly with the animal rights problem in the United States, but the same sort of campaigning has been going on for much longer in Britain, has been more intense and far more violent. There is hardly a university or pharmaceutical company in Britain that has not suffered from the attentions of the animal rights movement. There is every indication that the level of protest will continue to rise, until sufficient effort is devoted to telling the public the truth. The growing danger to animal research comes not only from the activity of the animal rights movement, but also from the inactivity of those who should be defending it.

MARK MATFIELD

*Research Defence Society,
Grosvenor Gardens,
London SW1W 0BS, UK*

SIR — Barbara Culliton wonders at the growing support for animal rightists in the United States. Part of the reason is that thoughtful members of the public troubled by — say — the use of chimpanzees in AIDS research can always hope that some sensible synthesis can be made of the animal rightist nonsense and medical establishment nonsense (of which the article in question is an excellent example). You do not need to be a philosopher or a theologian to know that there is something very dodgy about sacrificing chimpanzees for purely human purposes, and to know that this version of Cain's question ("Am I my brother's keeper?") — and others like it — requires a great deal of serious thought and public debate.

This is a vexed and troubling question to which I for one certainly do not know the answer. What I do know is that the debate that might help us find it is unlikely to be inaugurated by statements like "The animal rights people go for the heart, the biologists for the head". This is not a contribution to any debate, but an attempt to disqualify one of the possible parties to it.

If my teaching experience in undergraduate ethics is anything to go by, many people are ready for a much more serious discussion of these questions than the 'research community' is willing to offer. If it really intends to maintain this tone, it will lose the debate by default, and will deserve to. Someone has to make us think about these questions, and if it isn't the research community or the churches or the universities it will be the animal rightists: unattractive, bigoted and disingenuous though they may be. Cromwell would have approved "...since the work must go on, better plain men than none".

T. E. FORSTER

*Department of Pure Mathematics
and Mathematical Statistics,
University of Cambridge,
Cambridge CB1 2SB, UK.*

Astronomers still thinking too big?

SIR — One must give the advocates of "big science" credit for persistence in the face of adversity. One might have thought that the Hubble Space Telescope debacle would have taught US astronomers the folly of "putting all of your eggs in one basket", particularly if you plan to put that "basket" into high Earth orbit, where correcting errors will be both difficult and costly. Instead, at the top of the Bahcall committee's "wish list" for US astronomy, we find the Space Infrared Telescope Facility (SIRTF).

It takes only a simple exercise in arithmetic to see how lopsided the Bahcall committee's recommendations really are. Of \$1,550 million recommended for "large projects", SIRTF would get \$1,300 million. That \$1,300 million is larger than the total of \$1,222 million allocated to all moderate programmes and more than five times the \$251 million left for small programmes. Although the priorities of the Bahcall committee are presented as those of a limited astronomy community, I suspect that many of the astronomers supported by "small programmes" might have dissented, had their opinion been asked.

ROBERT J. YAES

*University of Kentucky, Medical Center,
Department of Radiation Medicine,
Lexington,
Kentucky 405036-0084, USA*

Digital display

SIR — In your article "Secret Service as ultimate referee" (*Nature* 350 553; 1991), you discuss the possibility that a person might transcribe data with a preference for certain digits. This was well documented nearly 40 years ago by H. Gysel (*Mikrochimica Acta* 267; 1953), in the case of analysts weighing samples for carbon and hydrogen elemental analyses where the last place of decimals of the weight had to be estimated. He found that there was an unconscious preference for certain digits that was fairly constant for each chemist, and the errors thus induced could influence the results by up to 0.25 per cent in the analysis. At the time, many journals required accuracy to within 0.3 per cent. Gysel discussed the allowance that should be made for this preference, and later (*Mikrochimica Acta* 577; 1956) he shows how replacing the balances with those using a different system helped.

This not only supports your argument, but it even suggests that if there is not a preference, your random number idea has in fact been used.

ALAN F. THOMAS

*Research Department,
Firmenich SA,
Case Postale 239,
CH-1211 Geneva 8,
Switzerland*

Diversity overrated

SIR — Once again, the spectre of genetic depauperization has reared its head, this time overriding the possibility that introduced disease may threaten the last wild population of golden lion tamarins¹. The News item in *Nature* reported that "plague is a risk we [the US National Zoo] feel we have to take to "maintain maximum genetic diversity". However, it is not clear that such a decision is justified.

There is general agreement that smaller populations are at greater risk of extinction than are larger populations^{2,3}. However, there is also increasing understanding that this risk is more likely to come from demographic constraints than from genetic constraints³. Although there is considerable evidence that environmental factors, including epizootic diseases, have had catastrophic effects on wild populations^{2,4,5}, there is a complete absence of examples of wild populations in which either inbreeding depression or a lack of genetic variability is known to have been primarily responsible for significant reductions in population size, much less extinction. On the contrary, there are numerous examples of wild populations that are both monomorphic and apparently healthy².

The continued disproportionate attention given to genetic considerations is due in part to the considerable monetary and public relations investment that many zoos have made in presenting themselves as conservation centres. Evidence of their success in these efforts is provided by the 'News' article, where it is stated that "most conservation programmes are . . . funded by zoos", which is simply untrue. It is also misleading to suggest that most zoos carry out re-introduction programmes. Rather, the goal of most zoos' captive breeding programmes appears to be the maintenance of captive populations, and only a handful have attempted to reintroduce endangered species into the wild.

All else being equal, there can be little doubt that genetic diversity is desirable. However, all is not equal, and poorly documented concerns about genetic diversity consume limited conservation resources and may also more directly threaten manipulated wild populations through disease transmission, outbreeding depression^{2,3} and translocation risks.

TRUMAN P. YOUNG

Center for Population Biology and
Department of Botany,
University of California,
Davis 95616, USA

1. Anderson, C. *Nature* **351**, 89 (1991).
2. Simberloff, D. A. *Rev. Ecol. Syst.* **19**, 473-512 (1988).
3. Lande, R. *Science* **241**, 1455-1460 (1988).
4. Harwood, J. & Hall, A. *Trends Ecol. Evol.* **5**, 254-257 (1990).
5. May, R. M. *Nature* **320**, 13-14 (1986).

War reparations

SIR — United Nations' resolutions require Iraq to pay 30 per cent of its future oil revenues to compensate the countries affected by the recent invasion of Kuwait, but this does not cover compensation for the damage to nature. We suggest that a sum of money be guaranteed from Iraq's oil revenues to be used to turn part of the wastelands we saw on television into a scene of a green desert. The chlorophyll of the trees would turn the carbon dioxide added to the atmosphere by the burning oil wells into wood and green leaves.

There is groundwater in the layers underlying the sand dunes of the desert of Kuwait and southern Iraq^{1,2}. The annual average rainfall is about 100 mm. A certain percentage of the rainwater infiltrates into the porous sand, to form a shallow water table. Below this, mostly fresh, groundwater, there is brackish and saline water.

We do not think there is enough water for intensive irrigated agriculture, but we are positive that with the available quantity of rain falling on sands, the area can be turned into a 'savanna' of desert trees and bushes. The tree roots can catch part of the water before it evaporates from the surface or penetrates into the groundwater. The preliminary results from our research on the hydrology of forestation in arid zones, indicates that sand dunes (under similar climatic conditions to that of the Persian Gulf) can retain a planted forest of tamarisk trees. Moreover, we found that in the tree-covered area, more humidity is retained in the upper soil than in the area bare of trees. The roots go down to the depth of 10 m and reach the brackish groundwater, thus ensuring themselves with a safe water supply, even during years of drought.

ARIE S. ISSAR
EILON ADAR

The Jacob Blaustein Institute for
Desert Research,
Ben Gurion University of the Negev,
Sede Boker Campus,
84990, Israel

1. Burdon, D. T. *Report to the Government of Kuwait* (Kuwait, 1966).
2. Himidia, I. H. & El Yaqubi, A. S. *3rd Seminar on Environment and Pollution* (Kuwait University, Kuwait, 1977).

Unwise after the events

SIR — Your publication of the correspondence from E. Guthy (*Nature* **348**, 670; 1990) and Frits Meijler (*Nature* **350**, 268; 1991) and recently by Falk Koenemann (*Nature* **350**, 648; 1991) and Rudolf Vrba (*Nature* **350**, 648; 1991) is at the same time startling and perplexing.

I recognize the broad readership of your journal and its interest beyond mere scientific publications. It surprises me nevertheless that your journal provides space for a

discussion on "German genes", on "national and racial characteristics" based on German history and for somewhat apologetic responses from German scientists born after 1950.

After the disastrous experiences of the past, is there really a need for a barely masked new racial view even when it originates from most tragic personal suffering and, though emotionally understandable, now inculcates subsequent generations?

What is more astonishing than the views of the individual authors is an editorial policy that provides a forum for discussions which are scientifically absurd and, in a unifying Europe, at best suited only to stir up emotions.

It might be a good idea to apply to such letters some of the principles supposed to be used for scientific publications.

HARALD ZUR HAUSEN

Deutsches Krebsforschungszentrum,
Im Neuenheimer Feld 280,
6900 Heidelberg 1,
Germany

SIR — I read with interest and irony the debate in your columns on the new Germany. But the debate seems futile.

Certainly we could ask such questions about many countries . . . Is there a new United States or is it the same country that practised genocide on its Indians, was the first and only country to use nuclear bombs in war, and practised dioxin warfare in Vietnam . . . ? Is there a new United Kingdom or the same one that practised genocide on Australian Aborigines, invented concentration camps in South Africa, and carried out saturation bombing with phosphorus bombs for three days in a row in Hamburg and Dresden . . . ? Is there a new Belgium since our conquest of the Congo? Do you want me to continue? In a civilized world one does not condemn the children for the crimes of their fathers.

J. E. DUMONT

Universite Libre de Bruxelles,
Campus Hopital Erasme,
Route de Lennik 808,
1070 Bruxelles,
Belgium

Thinking on the run

SIR — The term 'run' has come to connote an analytical session on a given instrument, conducted over a period of time. We have questioned several of our colleagues, and nobody knows the derivation of this now well established vernacular word. Does anybody know?

RONALD I. DORN

Geography Department,

JAMES CLARK

Chemistry Department,
Arizona State University,
Tempe, Arizona 85287-0104,
USA

The case for the human genome

A meeting next month in London should give a clear view of where the Human Genome Project is now, and where it is likely to be heading in the years that follow. The project (which is several) continues to gather momentum. The strategic questions have been settled, but there will be endless arguments on tactics. And ethical issues, while insubstantial, must be handled delicately.

FAMILIARITY exorcizes fear. That seems to be one of the unremarkable lessons learned so far from the great Human Genome Project or, more accurately, from the dozens of national projects that have sprung up in the past few years in different places.

Fears that research in biology would in future be conducted on automated assembly lines, or that research laboratories would be robbed by the demands of the sequencing plants of a whole generation of skilled technicians, have not materialized and are unlikely to do so, perhaps because none of the projects has yet entered what might be called its production phase.

Yet on the first issue, it is curious that Professor Walter Gilbert's impassioned statement of the reasons why biologists must learn to ask interesting questions of databases (*Nature* 349, 99; 1991) should have engendered so little irritation; perhaps biologists of all kinds look forward to the time when interesting questions can be answered. And the fear that technical skill would be monopolized by sequencing factories has been belittled by the quite remarkable continuation of the techniques of molecular biology even in the years since the campaign for a com-

plete sequence began.

Fears that human genome projects everywhere would rob other worthwhile projects of funds have similarly not yet bitten into the pattern of research supported by the grant-making agencies. Moreover, when people are talking realistically of automatic sequencing at a cost of no more than \$0.50 a base (with overheads included), it seems unlikely that a ten-year project to provide the complete sequence of 23 stretches of DNA, with 3,000 million bases altogether, should much exceed the annual cost of the present development phase.

The ethical objections to the sequencing of the human genome are necessarily more subtle. They range from the assertion that it would be improper that knowledge won in a scientific project such as this should be used to discriminate in novel ways between people seeking insurance cover or jobs, to the generalized alarm occasioned by any suggestion that eugenic improvement may be feasible, most obviously typified by the state of public opinion in Germany.

It is certainly the case that the complete sequence of a presumably representative human genome will be a more sure pointer

than any now available to the places in the human genome at which genetic abnormalities occasion genetic disease or a high frequency of somatic genetic change (as in cancer) — that is one of the important objectives of the exercise, after all. But the objection, if valid, also applies to what is already possible by the use of genetic markers for identifying particular variants of a genetic polymorphism.

In present practice in, for example, the insurance of persons, the trouble (and cost) that companies take to discriminate between their would-be customers is measured by the risks they are asked to undertake; it remains to be seen whether present ambitions to automate the sequencing process will be followed by a machine that can print out a detailed comparison between an arbitrary and some standard genome — and what fraction of, say, \$1,500 million the cost will be.

The reality of the use of a detailed knowledge of the human genome in discrimination between people is therefore almost certainly more distant than the fear. And the contention that people unlucky enough to carry identifiable genetic abnormalities should not be denied insurance on the same terms as other people begs the question why people relatively free from identifiable genetic abnormality should therefore pay more than would otherwise be necessary. In any case, the principle of discrimination has already been conceded both in respect of those who smoke cigarettes and, for much longer, those with a familial history of heart disease.

The eugenic argument is at once more shadowy and more special. The politically correct position, often defined without prompting by molecular geneticists, is that the human genome must not as such be interfered with, and *should* not be interfered with directly, even if that could be done safely (or without substantial risk of introducing genetic mutations). The chance that the precondition will be quickly satisfied in the direct manipulation of the human genome is so small that this fear is also more distant than it may appear.

But even if (some would say "when") safe manipulation were to become possible, it would be even more difficult than it is now to know on what ethical principle prohibition could be based. Known genetic defects can be (and are) already avoided in many countries by amniocentesis followed by abortion; if it were possible, would it not be ethically preferable that known defects should be

Playing the numbers game

Why is the genome project difficult? The numbers are a sufficient explanation. There are 3,000 million nucleotides in a single strand of the DNA distributed between the 23 human distinct chromosomes, which carry at least 30,000 genes (perhaps three times as many) comprising perhaps an average of 1,000 nucleotides each.

Not so long ago, it would have been a respectable outcome of a graduate studentship that a person should be able to isolate, clone (or magnify the amount of) a single gene and then determine its nucleotide sequence. That, it will be noticed, implies at least 100,000 man-years of effort merely to sequence the functional genes (taking three years as the average length of a PhD course).

Determining the sequence of the non-functional stretches of DNA is numerically a still greater problem — 30 to 100 times as great — but also one that could not conceivably be accomplished while research efforts are guided, as they must be, by tangible goals. And it would be improper to set armies of graduate students on the sequencing of long stretches of featureless DNA when there

might be nothing for them to interpret.

The second argument would be a sufficient justification for the attention given in the past five years to the automation of the sequencing process, but even so, the arithmetic of the genome projects remains formidable. Professor Isao Endo and his colleagues who (on page 89 of this issue), describe the Japanese development of an automatic sequencer, note that even a machine yielding 100,000 nucleotide bases a day would yield only 1 per cent of the entire human genome in a year, or the complete sequence of the human genome in a century.

Those who believe that the sequence of the entire genome should be the goal are forgiveably looking for still more productive automatic sequencing machines. Machines capable of sequencing one megabase of DNA a day would plainly make an important difference, bringing sequencing projects within the human work-span, for example. But it goes without saying that, with such sequencing performance, organizational matters — sample preparation and data handling and storage — will be the tails that wag the dog.

avoided for all time (or until the next spontaneous mutation) at the level of the germ line?

Practical eugenic questions are in any case not technical, but social. If the objective is to improve the capability of some human population, there is ample social research to show that substantial improvement of individual performance can be brought about by, for example, improved diet or by paying attention to the conditions under which children are brought up and educated. But much extant legislation on abortion already allows for the destruction of fetuses carrying unambiguously disadvantageous genes or genetic combinations.

So will not the culmination of the human genome projects be such an understanding of the relationship between individual fitness and genetic constitution that still more effective screening against genetic disadvantage will be possible and will become routine? The implied assumption begs all kinds of unanswered questions — the frequency with which heterozygotes for ostensibly disadvantageous genes are nevertheless advantaged (as in the classic case of sickle-cell anaemia); the frequency with which definable attributes of an individual's fitness are determined by multigenic systems as well as the frequency with which particular genes contribute to several attributes of the phenotype. (The dictum "One gene, one enzyme" does not imply "One gene, one recognizable attribute".) There is also the more practical question of women's willingness to suffer serial abortions in the pursuit of perfection in their children.

Merely to state these opinions is not to ensure that they will be generally accepted. Far from it. There is every likelihood that the general opinion of genome sequencing will remain cautious (to say the least) for a long time. That is one reason why the scientific community concerned owes it to itself to ensure that these questions are fully discussed. So much seems to be acknowledged: at the beginning of the US project, there was talk of setting aside a fixed proportion of the budget for the study of ethical issues. But guaranteed financial support is not nearly as necessary as general discussion within the community, accompanied by a recognition that questions that seem cut and dried to professional people may be deeply worrying for the more general public.

It is fair to guess that some quick success — perhaps the sure identification of hitherto cryptic but important genes responsible for disease — will be a boon in winning around the general opinion of the project to sequence the human genome. The danger is that intemperate certainty will have the opposite effect. It would be a shame if biology's megaproject, with all its promise, encountered the incoherent opposition that have blocked the path of great endeavours in other fields.

German (or, at least, what used to be West German) sensitivity about genetic manipulation appears similarly to spring from

recollections of social events, not of developments in genetics. Those who in the 1930s mounted the Holocaust hardly bothered to excuse themselves by reference to genetics, although the ill-defined and empty doctrine of Aryan superiority may have stilled the consciences of some practitioners. Is it reasonable, now, that the constructive promise of genetics should be diminished and neglected on such flimsy grounds?

For there is no doubt of the promise, and of the degree to which it has captivated the imaginations of those working in the human genome projects. It is unlikely that the practical benefits now foreseen of a complete sequence of the human genome can be an exclusive list; it may exclude some that eventually prove to be the most important. But even a listing of functional genes, like a listing of the components of a rifle (as in most armies' basic training for soldiers) would be of value; it would be interesting to know, for example, whether there are 30,000 or three times as many.

The rapid and sure identification of genetic diseases by comparison between DNA sequences from the tissues of an affected person and some reference sequence in a databank is the most obvious benefit and, given the high cost of projects in classical genetics, probably well worth an initial investment of, say, \$1,500 million. But means of telling whether somatic cells in disease differ from some standard genomic pattern may be much more important, not merely in the understanding of particular kinds of cancer, of psychiatric diseases and even of the normal degenerative processes of ageing. Understanding does not always presage more effective treatment, but the promise is high.

So, too, is the promise that the complete sequence of the human genome will yield an understanding of how the human genome itself has evolved. For this purpose, of course, only the complete sequence (rather than the sequences of only the possible functional genes) will suffice. But already enough is known of the interspecific relationships of genes with similar functions, and of the occurrence within an organism of families of genes that function in predetermined sequence (as with the pattern-forming genes of *Drosophila*, for example) to suggest that even sophisticated inspection of the database will yield important insights.

These two prizes, general though they are, would be sufficient to account for the enthusiasm that abounds, but there is more than that to say. The peculiar feature of the human genome project is that, for molecular biologists, it is at once a definable, an attainable and a majestic goal. It has also proved to be, in the past few years, a powerful stimulus of technique, not simply by the automation of what were previously labour-intensive operations, but also of novel ways of exploiting familiar procedures. (There is at present a spate of interest in modifications of the

polymerase chain reaction, or PCR, designed specifically to yield nucleic acid samples more suitable for use in sequencing projects that are certain to be invaluable in other contexts.)

Gilbert is probably right to argue that the end-result will be a sea-change in the way in which biology is practised. Quite what the change will be remains to be seen. It will probably be more profound than even Gilbert guesses.

One of the remarkable features of the genome project is that it has from the outset been international. Although much of the running is necessarily made by molecular biologists in the United States and by the two grant-making agencies chiefly involved (the Department of Energy through its national laboratories at Los Alamos, Argonne and Livermore as well as the National Institutes of Health), groups elsewhere have had a vivid interest in the description of the entire human genome for almost as long.

As an international enterprise, the genome project will pass yet another milestone at the meeting in London next month (18–22 August) which is known to the practitioners as HGMXI. The organizers, the international Human Genome Organization or HUGO, expect that 700 people will attend what is still called a "workshop", the latest in a series begun in 1973. The function of the meeting, like its predecessors, is severely practical — to add to the emerging map of the human genome, and to reach some kind of understanding of where efforts might next be concentrated. (No doubt the hidden agenda is the familiar business of learning, from what people say about their work, who does what well — and thus to identify opportunities for further collaboration.)

The hope now is that it will be possible to arrange for a change of style. A Howard Hughes Medical Institute grant to the Welch Memorial Library at Johns Hopkins University has made possible the development of a database (called the Genome Data Base), the first version of which was successfully tried out at a gathering last year at Oxford. The principle is that the database should contain information about securely mapped genes. For the time being, a securely mapped gene is one whose position on a chromosome has been established by linkage, and which is accompanied by the nucleotide sequence of the DNA to which the marker corresponds. By last September, there were 1,867 nuclear genes and 4,859 other DNA segments of unknown function in the physical map. The numbers of the participants expected at the London meeting point to one striking feature of the project — that it is in no sense a single project, but a necessarily loose confederation of many, some national (such as the National Center for Human Genome Research at the US National Institutes of Health), some international (such as the European Communities' Human Genome Analysis Programme) and some still based

on individual laboratories and research institutes.

HUGO itself owes its creation (at Geneva three years ago) at least in part to the wish of many people to allay fears that the large project then seeking funds from grant-making agencies in the United States would become a monolithic enterprise, innocent of contributions from elsewhere in the world and, perhaps, eventually in competition with similar enterprises elsewhere. But the organization was self-consciously intended as a means by which the pattern of development could be guided by working researchers and not by those, politicians and others, who administer national research funds.

But HUGO has no substantial funds with which to back research — perhaps the phrase should be “research and production” — so is an obvious target for the criticism that it is ineffectual. In reality, it can fairly claim a number of important successes; among other things, for example, it has been the umbrella beneath which the Paris-based French network CEPH (for Centre d’Etudes Polymorphisme Humaine) has become the prime source of uniform cellular material for national and international projects everywhere, thus reducing the uncertainty in knowing which genome is being sequenced.

It is also by now plain that the problems of international organization are likely to be more serious than were at first foreseen. For one thing, it is inevitable that a project on such a scale will engender chauvinistic ambitions — to be first with the complete sequence of a chromosome, to have developed the fastest automatic sequencer (for the time being) and so on.

There are also questions not yet accurately formulated about the relationship there will eventually be between the databanks originally established for storing all sequence information and those now likely to emerge from the mapping projects. Will the two systems eventually be merged, or will they march along in parallel with each other?

Even on a regional scale, the problems of coordination are formidable, and some would say unsolved. Dr Lennart Philson’s plea (*Nature* 351, 91–92; 1991) that coordination of molecular biology within Europe should be rationalized is heart-felt, but is unlikely to be quickly met. The strategic goal at which HUGO should be aiming is that of providing working researchers with such a secure sense that it can meet their felt need of closer collaboration that most will choose to shelter beneath its umbrella.

Finally, of course, it is inevitable that governments (and their electors) will eventually insist on having a say on important questions such as access to the data gathered. Fears that genetic data can be used for discriminatory purposes, however successfully dealt with in the years ahead, will never be entirely stilled. There are also difficult questions about the terms on which commercial organizations, perhaps pharmaceutical companies bent on drug design, should have access to

data gathered with public support. Conventions on this issue differ from one jurisdiction to another.

Despite its modest successes so far, HUGO will find it has to keep running hard if it is successfully to play the ambitious role it has set for itself. The long-term objective is to command the respect of the world’s genome sequencers individually. The best hope must be that governments will soon recognize the value of HUGO’s independence by providing the modest sums its modest coordination role requires. So far, the Soviet Union is the only government or public grant-making agency to have recognized the need (see page 3, this issue).

Even those who make light of the difficulties that lie ahead, believing them to be either tractable or likely to melt away in the face of further technical development, acknowledge that there must be a plan. What should it be?

The cornerstone of the US project is that the first five years should be spent in providing support for hopeful technical projects, in the belief that it will then be possible to make the most economical decisions about the means to use in the long haul. It is as if the designers of the Apollo Project (to put a man on the Moon) had, in the 1960s, begun by giving themselves the luxury of asking what kinds of rockets would best serve their purpose rather than by settling for the Saturn rocket already on the drawing boards. This decision by the US genome projects is not a mark of hesitation, but of enlightened deliberation, a way of seeking the most economical way of using rapidly evolving tech-

nology.

Meanwhile, attention will be concentrated on the construction of maps of the human genome — both linkage maps (based on classical genetic linkage studies, which provide a measure of the distance between genes from an assessment of the likelihood that they are inherited together) and physical maps (in which the positions of known genes and other DNA sequences are marked physically by a known and recognizable sequence of DNA, ideally distinct from other markers).

There are also to be sequencing studies of a series of model organisms, the old workhorse *E. coli* included. (It is strange that, when the argument about the setting up of the European Molecular Laboratory, as distinct from the European Molecular Biology Organization, was raging in Europe in the early 1970s, there were many who argued that sequencing of the K12 strain would be a project that would give the laboratory pointedness.) It remains the case that only 40 per cent of the *E. coli* genome is yet known, and that there are hopes of a more rapid and pointed harvest of information from the projects now mounted for the study of still simpler organisms.

Of these, potentially one of the most fruitful is the scheme for producing the complete sequence of the nematode *Caenorhabditis elegans* by means of a collaboration between the British Medical Research Council’s Laboratory of Molecular Biology at Cambridge and the Medical School of the Washington University, St. Louis.

Fishing for genes

Enthusiasm for the use of cDNA as a substrate for the sequencing of the human genome, and the potential promise of this route, is amply illustrated by the recent work of J. Craig Venter and his colleagues, of the National Institute of Neurological Diseases and Stroke, at the National Institutes of Health (*Science* 1,651; 1991).

Using cDNA from human brain tissues (hippocampus, temporal cortex and fetal brain), Venter’s group used an automatic DNA sequencer to make copies of cDNA sequences selected more or less at random from a commercial cDNA library and used an automatic sequencer to generate copies of them corresponding to the functional parts of genes, and which are called “expressed sequence tags”.

In the event, there were 600 useable pieces of DNA of this kind, on the average a little less than 400 bases long, whose sequences were determined automatically. The same materials were then used for seeking matches in the nucleotide and protein sequence databanks.

The remarkable finding is that no fewer than 337 of the cDNA sequences represent “new” genes in the sense of being coding sequences of DNA whose con-

stitution as human genes has not previously been recognized. But 48 of the products were recognized as similar to genes occurring in other organisms — there is a *C. elegans* collagen in human brain, for example.

Venter and his colleagues are mostly at pains to use their experiment as a means of making their case that cDNA libraries are an efficient approach to the listing of all the functional human genes. They also provide some evidence that it should be possible to reduce the frequency with which common genes turn up in their samples by screening one cDNA library against another.

But they acknowledge, as they must, that the technique does not provide information about the positions of genes in the genome (although it can be used, with sufficient material and a battery of hybrid cells containing different sub-sets of human chromosomes to tell which chromosomes carry the genes concerned.) They argue, in response to the assumption of the end-to-end sequencers that it should be easy to tell a gene from a piece of nonsense DNA that this is not necessarily the case when the pieces concerned are very small.

Watches make machines

The Seiko Company is best known not just for the watches it makes, but for having pioneered their automatic assembly. In its research and development building on the outskirts of Tokyo, there is an automatic watch factory whose function is simply to train those who will manage automated watch factories elsewhere, at secondary manufacturing centres such as Singapore, for example.

The notion that Seiko should involve itself in the design of an automatic DNA sequencing machine goes back ten years, and is the concept of Professor A. Wada, then a professor of physics at the University of Tokyo and the chairman of a committee set up to worry about instrumentation for biotechnology.

The outcome of the project is the automatic sequencer now installed at the RIKEN laboratory (*Nature* **351**, 593; 1991). But much of the interest of the project lies in the way in which it has been carried through. On three separate visits between 1984 and 1989, it seemed that the same cheerful group of engineers had reached a new plateau of cleverness.

The robot arm designed to select DNA

fragments was early in evidence, as was the electrophoresis tunnel. At the outset, Seiko's collaboration with the Fuji Film Company was based on the assumption that the analysis of gels would replicate the standard laboratory technique of exposing X-ray film to radionuclides, but the technique was abandoned as much because the prospect of using fluorescent tags for nucleotides that could be interrogated by lasers is plainly better suited to automation.

Throughout this period, just four engineers (with workshop backup) seem to have been responsible for the development. At no time have they pretended to be molecular biologists. From the outset, they have made it plain that their task is an exercise in ingenuity that will not be complete until the separate automatic functions of the process are thoroughly integrated into a truly automated plant.

The most striking feature of the development has been its informality. The small group involved always seemed entirely self-sufficient. What they are doing now is not recorded, but it is unlikely that they will have gone back to making watches.

The potential prize turns on the relatively modest size of the complete genome — at 100 million bases, roughly the size of a human chromosome — and the circumstances that the genetic pattern of the organism is splendidly hard-wired; there are 956 cells only in the complete organism — fewer cells than there are genes — and that the lineage of each cell has now been determined. The *C. elegans* project (which should yield a physical map in two years and the complete sequence soon afterwards) is thus a rehearsal not merely for the technology of sequencing, but for the intellectual problems that will have to be faced when substantial tracts of the human genome are available.

Each of the other model organisms occupying the genome projects has its own advantages. Yeast is easily manipulated and is important in biotechnology, the cress called *Arabidopsis thaliana* has the prime advantage of being a plant with a five-day life-cycle and chromosomes roughly a tenth as big as those of mammals, while *Drosophila* and the mouse have in their favour decades of careful classical and, lately, molecular genetics. (Given that the known corresponding genes of mice and men are similar to the extent of 70 per cent, it might not be unreasonable to regard a complete sequence of the mouse genome as a reasonable substitute for that of the human genome.)

One certain development in the next few years is that the list of model organisms will quickly grow as organizations with differing interests prepare the ground for more ambitious projects. Among plants, maize and rice are obvious candidates, as are the parasites of tropical disease.

Meanwhile, there has emerged an informal understanding that the physical mapping of the human chromosomes should be parcelled out among various laboratories. The immediate objective is the construction of a physical map of the underlying structure, which in the context of the US programme is a well-specified ideal — the identification of nucleotide sequences at intervals of about 100,000 bases along the DNA structure which can then be used as primers for a polymerase chain reaction (PCR) that will yield essentially indefinite amounts of the intervening stretch of DNA.

Simple arithmetic suggests that there will need to be 1,000 markers for each human chromosome, which is far from a trivial requirement. Yet this is the easiest way of constructing a physical map of a DNA structure, when the length of intervening stretches of DNA (measured as the number of bases) will ultimately provide a way of telling where genes are — provided that the mapping process is not too seriously impeded by long stretches of featureless DNA. (If there are problems of that kind, they should quickly become apparent.)

For sequencers, there are obvious benefits in such physical maps. For example, it would then be possible to deal with unambiguously defined stretches of a genome, relying on known subsequences a few dozens of nucleotides long which, moreover, can always be resynthesized and which are thus replicable milestones in otherwise unknown territory. To chromosome mappers, this is the equivalent of providing objectively measurable geodetic coordinates to a terrestrial navigator.

The five-year hiatus in the US programme, admirable though its deliberation may be, is nevertheless a frustration, both in the United States and elsewhere. That may account for the recent groundswell of opinion, especially in Europe, that there should be a more direct attack on the identification and sequencing of the function genes in the human genome. There are several incentives: many are anxious that there should be some progress while they are still personally active in the field, others argue that the generous support that will be needed at later stages of the projects will be more readily won if there were something of outstanding value to show for what has been done already.

Perhaps this explains the recent spate of enthusiasm for the notion that the place to start is with the sequencing of complementary DNA (cDNA) — that obtained from messenger RNA (mRNA) in real cells by the use of reverse transcriptase to convert RNA to DNA. In principle, at least, a sufficiently thorough sampling of the tissues of an organism should yield a complete set of functional genes, while the comparison between the cDNA libraries should give some indication of which genes are active only in some tissues.

This is one of the topics set aside for separate discussion at next month's HGM XI in London. There are technical questions of some importance — how to ensure that relatively inactive genes are fully represented in the material used for sequence determination, for example. But the underlying issue is political (as these things go): should there be a deliberate attempt to provide a foretaste (see box) of what the human genome projects will eventually provide? Much may hang on whatever consensus emerges from the meeting.

Even the enthusiasts for cDNA sequencing acknowledge that such procedures will reveal neither the way in which real genes are broken up by apparently irrelevant introns nor the nature and position of the relevant control elements. But the band of enthusiasts has recently been growing quickly with the recognition that a listing of all (or almost all) of the functional human genes would in itself be a worthwhile prize, perhaps even the most valuable prize to be won from complete sequencing of the human genome, with all its "junk DNA".

These, however, are almost technicalities. What is now clear is that the sequencing of a complete human (or other mammalian) genome is technically possible and that the benefits, while incalculable, will be immense. For the scientific community, perhaps the most cheerful feature is that, in what has so far been accomplished, amity and the prospect of true international collaboration has not been undermined by the substantial intellectual arguments engendered by the mounting of biology's most ambitious project yet. That is more than a little to be grateful for.

John Maddox

AIDS against the rest of the world

The spread of AIDS through heterosexual contact in Asia and Africa has reached truly epidemic proportions. The rest of the world should take notice while there is time.

THE most important news to come from the seventh international conference on AIDS that just ended in Florence is the worst news possible. While Western nations live with the illusion that the human immunodeficiency virus (HIV) has been somewhat contained within populations of homosexuals and intravenous drug users, data from the World Health Organisation make plain that the heterosexual spread of AIDS in Asia and Africa constitutes an epidemic that threatens the very fabric of the societies affected.

According to figures presented by James Chin of WHO, by the end of this decade 40 million men, women and children will be infected worldwide. In just three years time, he predicts, 10 million people in Africa will be infected; in Asia, where Thailand and India are particularly at risk, 3 million people are likely to become HIV positive. Most of those people will be young men and women in their 20s and 30s, whose deaths in huge numbers will devastate the population. WHO chief Hiroshi Nakajima says the looming pandemic threatens "the socio-economic development, and the very survival of whole communities and countries".

To those who closely follow the epidemiology of AIDS, these figures are not entirely a surprise. Nevertheless, the potential force of the pandemic has been brought home anew by the Florence meeting, where preliminary data on the biology of the AIDS virus makes it clear why a massive tragedy in Asia and Africa portends disaster in the West as well.

For the past decade — in fact, ever since AIDS was first reported in 1981 by the US Centers for Disease Control in Atlanta — Western researchers have emphasized the view that AIDS is difficult to transmit. And, indeed, the epidemic appears for now to have peaked in Western nations where HIV spread has been more-or-less contained among high-risk groups where the number of HIV positive individuals is likely to reach 2 million by the middle of this century. Such numbers are hardly a comfort. However, people have been assured that only unusual homosexual practices, such as anal intercourse, or the sharing of dirty needles among drug users really puts one at risk of getting AIDS. In many regards this is reassuring. Large numbers of men and women can go to sleep at night not worrying about contracting a lethal virus.

For the time being, this is largely true, but only because the virus is not yet as widespread among the heterosexual population as it is in Asia and Africa where the virus spreads with apparently fearsome ease from

man to woman and woman to man.

Studies suggest that the upward spiral of HIV infection in Thailand can be traced to the country's large number of prostitutes — government officials estimate there are 800,000 women working as prostitutes — and the fact that sex with prostitutes is common. One recently reported survey claimed that 75 per cent of Thai men said they have had sex with prostitutes.

Figures recently published by Jim McDowell, the US congressman/physician who is co-chair of the Congressional Task Force on International AIDS, also paint a bleak picture for India. According to McDermott, in Bombay alone, the HIV infection rate among 100,000–150,000 prostitutes has jumped from one per cent in 1987 to 30 per cent in 1990. With prostitutes averaging six contacts per night, and an HIV transmission rate of 0.1 per cent, he estimates that 6,000 men are being infected every month in Bombay alone.

The numbers of men from other nations who purchase what is now known to be high-risk sex in these and other countries can only be guessed at, but it is naive to think that once HIV is firmly embedded in the heterosexual population of international centres such as Bangkok and Bombay, its outward spread can then be contained.

Then there is the issue of HIV spread through what are now regarded as highly unlikely means — from physician to patients, for example, or vice versa. In the United States, Britain and the rest of Europe, current data seem to argue that such transmission is rare to the point of being nearly impossible.

There is a striking counter example — that of a dentist named David Acer in a medium-sized Florida town who, before his death last year, passed on HIV to at least one patient whose death is now being chronicled in the public press. Kimberly Bergalis, age 23, convincingly says that she has never had sex and never used drugs. Indeed, her only exposure to HIV appears to be at the hands of her dentist, who is presumed to have transmitted the virus in the course of an ordinary dental procedure.

Bergalis, who has written vividly about losing weight in her fingers and enduring the growth of fungus like 'white fur' all over her mouth and lips, has come to symbolize a fear that the medical profession has done its best to vanquish. Even if it is true that one is not likely to get AIDS from one's physician, the Bergalis case has created understandable public unease and generated a new fear of

this lethal disease that is entirely rational.

One thinks here of Camus's town of Oran and the day, sometime in '194-', that "When leaving his surgery on the morning of 16 April, Dr Bernard Rieux felt something soft under his foot" — namely a dead rat whose presence he first thought "rather odd, no more than that". But as rats in great numbers began dying, people finally recognised that the plague was upon them and that drastic measures had to be taken. It is not impossible that the few cases of unusual AIDS transmission carry a message we do not want to hear.

Controversial data presented in Florence by William Haseltine of the Harvard-affiliated Dana-Farber Cancer Institute in Boston speaks to a possible route of easy heterosexual transmission. Haseltine argues that dendritic Langerhans cells in the mucous lining the mouth and genital areas may be an important site of infection. Haseltine reports that once these cells are infected by HIV, they produce 80 per cent of HIV in the body — even more than white blood cells. In Haseltine's opinion, it is time to recognise AIDS as a "lethal venereal disease — primarily heterosexual".

Although many of Haseltine's colleagues challenge his views (while he, himself, notes that support for the work on dendritic cells comes not from the US National Institutes of Health but from private sources), the idea that AIDS is really a heterosexual disease is an idea whose time may have come.

Certainly, one's perspective on the disease is pertinent to the heated political debate in Florence that took place over US immigration policy and the next international conference scheduled for Boston in 1992. As it stands, US law forbids travel visas or immigration to persons known to be HIV positive. Taking the view that AIDS is not easily transmitted, and that a ban on travel is really a political decision to ban homosexuals rather than a decision about public health, scientific leaders and gay AIDS activists have threatened to cancel the 1992 conference unless the law is overturned. The issue is cast as one of moralistic conservatives versus compassionate liberals. Cast that way, there is only one side to be on.

But casting the debate that way totally ignores a more important reality that needs to be addressed in medical terms, not political. Is AIDS more easily transmitted than Western researchers have wanted to believe? Is the threat of a worldwide pandemic real and imminent? We need to know.

Barbara J. Culliton

Control engineering

Michael E. Hochberg and Jeffrey K. Waage

REPORTS on pages 82¹ and 85² of this issue take the development of effective 'biopesticides' a long step further. Tomalski and Miller¹ and Stewart and colleagues² show that the rate at which an insect virus kills a pest can be increased by incorporating toxin-producing genes from other arthropods into the viral genome.

Concern about the widespread use of

thuringiensis, where the problems of pathogen application and persistence in the field have been addressed by building the bacterium's toxin-producing genes into other, hardier bacteria or into plants themselves. Commercialization of these imaginative products is underway, but is now dogged by the appearance of resistance to the *B. thuringiensis* toxin in a range of insect pests

for long-lived viruses with fast kill rates. But there is a large body of theory³⁻⁶ on the population dynamics of pathogens which suggests otherwise (see Fig. 2). For instance, say the engineered pathogen is introduced to the pest population only once, or very infrequently, with the goal of long-lasting control. Theory^{3,4} on the dynamics of infectious disease would predict that two of the most important factors determining the ability of the pathogen to regulate the pest population to economically acceptable levels are the kill rate and the persistence of the pathogen in the environment. As kill rate is increased from very low levels (for example, through the incorporation of a relatively benign toxin-producing gene), there is first an increase in the ability of the pathogen to regulate and depress the pest population³.

But, as the ability of the engineered pathogen to kill the pest quickly is increased, population levels of the pest increase (because less pathogen is produced by each dying pest) and there is a propensity for the pest population to cycle (the period of which is inversely related to the environmental persistence⁴). Still further increases in pathogenicity (or decreases in persistence) are more likely to result in the extinction of the pathogen than the eradication of the pest itself. Even if the kill rate of the pathogen is adjusted to give adequate control of the pest, there is always the danger of the introduction resulting in the elimination of other specialist natural enemies of the pest from the system⁵. The continued presence of these other natural enemies, especially insect parasitoids, can supplement the control offered by the pathogen. Should the pathogen population itself die out, natural enemies can prevent the pest from resurging to outbreak levels.

Now envisage a situation in which the pathogen population is frequently augmented, such as would be the case for many biopesticides (Fig. 2). Here, population dynamics theory^{4,6} predicts that if the pathogen is added at low to intermediate rates, there will be a tendency towards more constant pest populations than would have been so if the pathogen were left to its own devices to control the pest. But it is only when augmentation rates approach a critical threshold (determined by the various biological parameters of the system) that pest numbers are substantially reduced, and there is a possibility of eradication. Again, as with pathogens that succeed in controlling the pest after only a single application, other specialist natural enemies of the pest are likely to suffer. The consequence of a reduction in the diversity of other natural enemies is that if the application of the pathogen ends before eradication of the pest, there is a risk

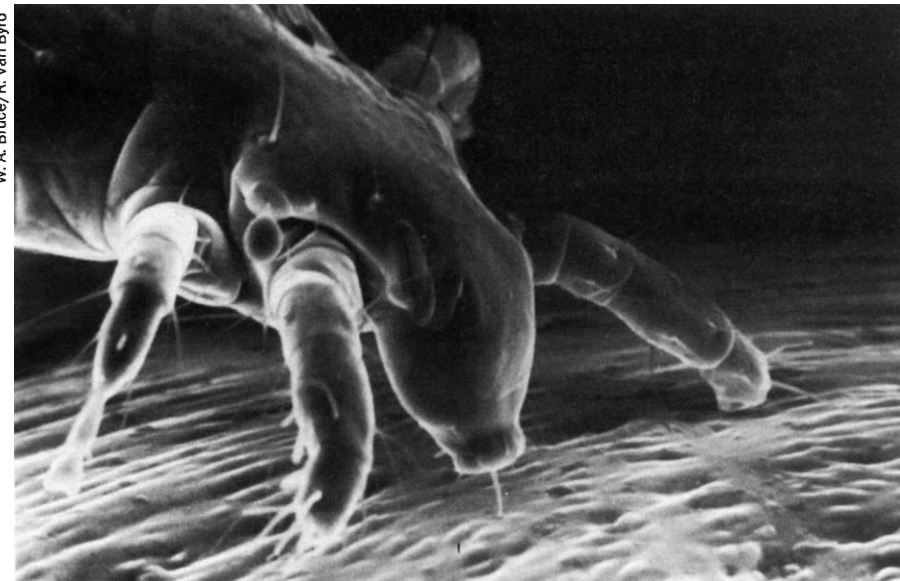


FIG. 1 Prey paralysis — a female of the mite species *Pyemotes tritici* feeds on an insect after paralyzing it with a toxin. It is the gene for this toxin which Tomalski and Miller¹ have engineered into an insect-specific baculovirus.

broad-spectrum, chemical insecticides has led to a surge of research into alternative pest control technologies. The formulation of insect pathogens as biopesticides has attracted particular attention because of their specificity to the pest and potential for commercial development. But biopesticides have yet to secure a sizeable slice of the pesticide market — globally, they account for only about 1 per cent of a sales volume of \$25,000 million, which in part reflects technical snags with their development. Formulated insect pathogens often break down quickly on plant surfaces; they may be costly to produce; and they kill pests more slowly than chemical insecticides.

It is this third drawback that the authors have tackled. They demonstrate that genetic fragments coding for the expression of natural toxins derived from the venom of mites¹ (Fig. 1) and scorpions² can be inserted into the genome of insect-specific baculoviruses and expressed within the cells of living insect hosts. The intracellular secretion of these toxins significantly reduces the life-span of the target insect (and potential damage to agricultural crops) as compared to infection by the non-engineered virus.

Most other work on the expression of foreign genes for pest control has been done on the insect-specific bacterium, *Bacillus*

— a lesson, perhaps, for those trying to engineer proteinaceous toxins into insect viruses.

In much of the new pest-control technology, engineered insect pathogens are viewed solely as sprayed-on products of short persistence. But a further bonus of the work by Tomalski and Miller and Stewart *et al.* is that they demonstrate that pathogens can also be engineered for long-term environmental persistence. Insect baculoviruses generally protect themselves from the vagaries of the environment through encapsulation in a polyhedrin protein coat. It just so happens that a single gene codes for the expression of the polyhedrin protein, and this gene can either be replaced or accompanied by the toxin-expressing gene. If the polyhedrin gene is replaced, progeny virus will quickly perish when liberated after the death of the pest. But if the ability of the virus to envelop itself in its protein coat is preserved, then the virus can be transmitted readily from pest to pest. The expression systems employed by Tomalski and Miller and by Stewart *et al.* are flexible enough to encourage or alternatively remove the ability of the virus to secrete its polyhedrin coat.

The genetic engineer therefore has a substantial amount of control over both pathogen persistence and kill rate. At first sight it would seem reasonable to engineer

1. Tomalski, M. D. & Miller, L. K. *Nature* **352**, 82-85 (1991).

2. Stewart, L. *et al.* *Nature* **352**, 85-88 (1991).

3. Anderson, R. M. *Nature* **279**, 150-152 (1979).

4. Anderson, R. M. & May, R. M. *Phil. Trans. R. Soc. Lond. B* **291**, 451-524 (1981).

5. Hochberg, M. E., Hassell, M. P. & May, R. M. *Am. Nat.* **135**, 74-94 (1990).

6. Anderson, R. M. *Parasitology* **84**, 3-33 (1982).

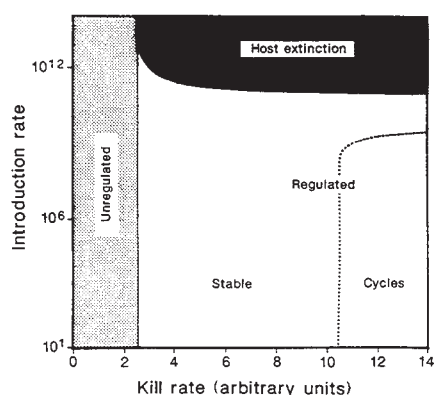


FIG. 2 The theoretical effects of pathogen kill rate and pathogen introduction rate on the population dynamics of an insect-host/viral-pathogen system. At low rates of introduction, the pathogen regulates the host to constant numbers only for intermediate kill rates. But as introduction rate is increased, intermediate to high kill rates tend to dampen oscillations in pest numbers. Only at very high introduction rates is pest eradication possible. (See ref. 6 for further details.)

that the pest will resurge to levels higher than those present before the pathogen population was augmented.

There are, too, broader implications, if the work described here reaches the stage of application, in that the prospect of engineered viruses persisting in the environment troubles some people. Will such viruses exclude natural forms or find new, undesirable niches? Can engineered genes escape from their viral vector? If so, can such genes successfully replicate and be expressed in

unintended hosts? Although there is no clear consensus on these and other risks involved in the release of genetically engineered baculoviruses, disabling engineered viruses so that they do not persist has some appeal. The highly restricted host ranges of most baculoviruses, and the ephemerality of those which are engineered not to express their protective protein coat, should give at most a small cause for concern. □

Michael E. Hochberg is at the Centre for Population Biology, Imperial College, Silwood Park, Ascot, Berkshire SL5 7PY, UK, and Jeffrey K. Waage at the International Institute of Biological Control, Silwood Park, Ascot, Berkshire SL5 7TA, UK.

MOLECULAR CHAPERONES

Unfolding protein folding

T. E. Creighton

'MOLECULAR chaperones', the cellular proteins that assist in the assembly of oligomeric proteins, are aptly named — the term conjures up visions of the human chaperone, whose job, put biochemically, was described by Ellis¹ as preventing "improper interactions between potentially complementary surfaces and to disrupt any improper liaisons that may occur". Indeed, molecular chaperones seem to act primarily by preventing partially folded and unassembled protein subunits from aggregating and precipitating. But how do they recognize a variety of different proteins and discriminate between proper and improper partners? How do they release the protein only when it is in the appropriate folded or assembled form, and why does this process require energy? And is the chaperone a more active 'match-maker', in encouraging the proper folding and association of the protein molecules? The outlines of answers to these points become clearer with the report of Martin *et al.* on page 36 of this issue².

The best-characterized molecular chaperone is GroE from bacteria, an abundant protein that is required for cell viability and for assembly of various other proteins^{3,4}. Similar molecules are found in mitochondria and chloroplasts, where they are required for the folding and assembly of at least some imported proteins⁵. GroE consists of two different components, GroEL and GroES (also known respectively as cpn60 and cpn10 or as hsp60 and hsp10); GroEL exists as a complex of 14 identical polypeptides (a

14-mer) of 549 residues each, whereas GroES is probably a 7-mer of chains with 97 residues each^{3,4}.

Studies of the effects of GroE on the refolding and assembly of dimeric^{6,7} and monomeric⁸ enzymes have demonstrated that GroEL binds the unfolded proteins tightly and releases them in a folded form, either assembled or able to become so, after hydrolysis of ATP and a physical interaction with GroES. Martin *et al.* have extended such studies considerably, using three unfolded proteins: casein, which is permanently unfolded; dihydrofolate reductase (DHFR), which refolds spontaneously and rapidly; and rhodanese, which normally aggregates rather than refolds. These proteins are monomeric, so only their refolding was involved, not assembly into higher forms.

Chaperones such as GroEL bind only unfolded proteins, not those that are fully folded. Although most unfolded proteins are thought to be in the form of random coils under denaturing conditions, a collapsed 'molten globule' type of conformation⁹ seems to be their natural state under more physiological conditions¹⁰ and the one most likely to be recognized by a chaperone. Nevertheless, unfolded proteins are very flexible and need not adopt the same conformation when bound to the chaperone.

Direct studies of the bound protein were necessary, and have been performed by Martin *et al.* Such studies are made difficult by the much greater size of GroEL than the

protein it binds; DHFR contains only 189 of the 7,875 residues present in its complex with GroEL. Fortunately, neither GroEL nor GroES contain tryptophan, so the fluorescence properties of the tryptophan residues in both DHFR and rhodanese could be used, as well as their protease-sensitive and dye-binding properties. Martin and colleagues found that the properties of bound DHFR and rhodanese were very similar, and like those expected if both were in the molten globule type of conformation. The unfolded proteins bound to GroEL were still sensitive to proteases, so they were not completely sequestered from the bulk solution.

Upon adding ATP, the bound DHFR quickly adopted physical properties closer to those of the native protein and was released from GroEL. The related DHFR and rhodanese were still not refolded, however; rhodanese tended to precipitate and DHFR tended to bind again. ATP was hydrolysed so long as unfolded protein molecules were bound to GroEL; 27 molecules of ATP were hydrolysed per molecule of rhodanese released. These results indicate that GroEL simply binds unfolded proteins in a compact form, and that ATP hydrolysis causes their dissociation.

Dihydrofolate reductase tended to refold spontaneously after release from GroEL, but refolding of rhodanese required the presence of GroES. GroES also stimulated refolding of GroEL-bound DHFR, and increased ATP hydrolysis to 130 ± 20 molecules for each molecule of rhodanese refolded. Whether GroES actively catalyses the folding process on GroEL, in terms of lowering the free energy of the transition state, or simply inhibits nonproductive processes, is not yet clear. One interpretation of the data is that the role of GroES is to keep the proteins bound, protecting them from aggregating with other unfolded protein molecules until they are fully refolded.

Martin *et al.*² went further and measured the stoichiometries of binding by GroEL. These have considerable implications. GroEL is a 14-mer, yet it binds productively only a single molecule of unfolded DHFR, rhodanese² or pre- β -lactamase⁸. The 14 GroEL monomers are arranged in two rings of seven, packed face-to-face⁴ (see figure). The only plausible position for a single binding site on such a symmetrical structure is at its centre⁸, where there is a hole. There should be 14 identical binding sites there, but only one might be occupied at any time owing to steric hindrance.

Such an arrangement is understandable if a protein is to refold while bound to a chaperone, for it should be bound tightly, but not so tightly as to inhibit refolding. The presence of 14 identical sites with individually weak affinities would cause a protein to be bound to any one of them more frequently than to an individual site. Such a loosely bound protein could still refold, for folded conformations would be more stable because

Reverse polarity

ONE of the characteristic features of autosomal dominant polycystic kidney disease, a common and fatal genetic disorder which maps to chromosome 16, is the accumulation of fluid in the renal tubules. Immunolocalization studies by P. D. Wilson *et al.* (*Am. J. Physiol.* **260**, F240–430; 1991) now show that the integral membrane protein Na⁺, K⁺-ATPase is confined to the apical (lumen-facing) membrane of the renal tubule epithelial cells of patients, whereas in normal cells it is found in the basolateral membrane. From analyses of cultured epithelial cells, it seems that Na⁺ ions are consequently pumped into the lumen instead of out of it, which would explain the large amounts of fluid secreted into the renal tubules.

Faint hope

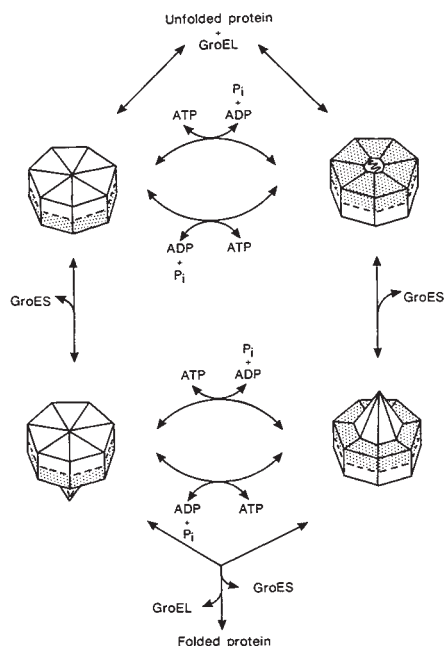
KNOWLEDGE of the ratio of deuterium to hydrogen created in the Big Bang would provide a strong pointer to the density of the early Universe. But the material seen in the stars nearest to us — even 'metal-poor' ones — is the product of billions of years of stellar processing and the original ratio may well have been distorted. The answer, according to J. K. Webb and colleagues (*Mon. Not. R. astr. Soc.* **250**, 657–665; 1991), may lie in the spectra of quasars. The light from these most distant of observable objects passes through clouds of relatively primitive gas, known as Lyman- α clouds after the hydrogen absorption lines seen in the spectra. Traces of deuterium in the clouds may also be just detectable and a reasonable measure of the ratio, confirming or excluding the locally reckoned values, should be possible.

Bloom with a view

THAT the biology of hummingbirds can be explained in terms of energetics runs counter to the general observation that animals trade time spent feeding against minimizing the risk of being eaten themselves; it is not that hummingbirds do not provide juicy bite-sized morsels, but that they are good at staying out of trouble. Experiments by S. L. Lima (*Evol. Ecol.* **5**, 220–230; 1991), using sugar-water feeders shaped like flowers with and without a corolla, suggest that female Anna's hummingbirds (*Calyptrorhynchus*) from Arizona are especially vigilant when feeding from large flowers, frequently withdrawing from blooms to survey the scene. They also prefer to feed high above the ground, in case they end up as fast food for roadrunners (*Geococcyx californianus*). Evolution has thus forced them into the enviable position of enjoying both the sweetest cocktails and the best views.

Model for the action of GroE in refolding proteins suggested by the data of Martin *et al.*². GroEL is depicted as a nonsymmetrical dimer of two rings of 7-mers. A single unfolded protein molecule can bind to any of seven sites in the central channel of only one of the rings (that stippled). The two possible isomers of the dimer are interconverted by a quaternary structure change¹⁴. This would require ATP hydrolysis when unfolded protein is bound, for the protein must be displaced from one 7-mer and transferred to the other. The unfolded protein would probably move through the channel between them, but it could also dissociate into the bulk solvent in the absence of GroES. The protein would tend to refold spontaneously while being transferred from one site to the other.

GroES is depicted as binding similarly to just one 7-mer of GroEL, not necessarily the same as that to which the protein is bound, and it must also be displaced from one to the other during the quaternary structure change. The function of GroES is postulated simply as preventing the bound protein from dissociating from GroEL until it has completed refolding, when it no longer binds to the sites on GroEL. Consequently, ATP hydrolysis ceases, GroES dissociates and the folded protein is released. This scheme does not include any steps for assembly of the refolding protein into higher order structures, which would have to occur in solution after release of the refolded monomers.



they have numerous interactions that occur simultaneously, not individually. A similar principle has been used to demonstrate that unfolded proteins can refold while bound reversibly to ion-exchange resins¹¹, which can be considered an artificial type of chaperone.

GroES is believed to exist as a ring of seven identical polypeptide chains^{3,4}, a structure that would be well-suited to bind to the open face of the GroEL disk, for both have seven-fold symmetry¹² (see figure). Two GroES 7-mers should bind to a symmetrical GroEL disk, yet Martin *et al.* show the functional stoichiometry to be one GroES 7-mer for each GroEL 14-mer. This suggests that GroEL is a nonsymmetric dimer of 7-mers. Moreover, GroES binds to GroEL only when ATP is being hydrolysed⁴ (that is, when the unfolded protein is being released and re-bound by GroEL). There is a coupling on GroEL between binding of unfolded protein, hydrolysis of ATP and interaction with GroES.

So the situation is a highly dynamic one, and it is tempting to propose that there is 'half-of-the-sites-reactivity'¹³ that alternates between the two halves of GroEL as a result of a quaternary structure change. The binding of both GroES and the unfolded protein would occur first to one and then to the other half of the GroEL 14-mer. The quaternary structure change would need to be driven by ATP hydrolysis and would cause the unfolded protein to be transferred from the centre of one 7-mer ring to the other, during which time it might refold spontaneously or dissociate. GroES would also bind alternately to only one of the two asymmetric halves of GroEL and would keep the unfolded protein sequestered from

the bulk solution. Once refolded, the protein would no longer bind to GroEL, ATP hydrolysis would stop, the GroES would dissociate, and the refolded protein could be released.

This scheme is speculative, but it has the appealing properties of being consistent with what is known about protein structure, function and folding *in vitro*, of not requiring an active role in folding of GroES, and of providing a plausible explanation for the unusual seven-fold symmetry of GroE. The model does not include a mechanism for assembling a multimeric protein, but this process could occur after release of the monomer into solution if the only function of GroE is to refold monomers by preventing them from precipitating when incompletely unfolded¹². Given the amount of attention now being paid to chaperones, there should not be too long to wait for a verdict on this or any other model. □

T. E. Creighton is in the European Molecular Biology Laboratory, Postfach 10.2209, Meyerhofstrasse 1, 6900 Heidelberg, Germany.

1. Ellis, R. J. & Hemmingsen, S. M. *Trends biochem. Sci.* **14**, 339–342 (1989).
2. Martin, J. *et al.* *Nature* **352**, 36–42 (1991).
3. Hemmingsen, S. M. *et al.* *Nature* **333**, 330–334 (1988).
4. Georgopoulos, C. & Ang, D. *Seminars Cell Biol.* **1**, 19–25 (1990).
5. Hartl, F.-U. & Neupert, W. *Science* **247**, 930–938 (1990).
6. Viitanen, P. V. *et al.* *Biochemistry* **29**, 5665–5671 (1990).
7. Buchner, J. *et al.* *Biochemistry* **30**, 1586–1591 (1991).
8. Laminet, A. A. *et al.* *EMBO J.* **9**, 2315–2319 (1990).
9. Kuwajima, K. *Proteins struct. funct. Genet.* **6**, 87–103 (1989).
10. Pitts, O. B. *et al.* *FEBS Lett.* **262**, 20–24 (1990).
11. Creighton, T. E. in *Protein Structure, Folding and Design* (ed. Oxender, D. L.) 249–257 (Liss, New York, 1986).
12. Gatenby, A. A., Viitanen, P. V. & Lorimer, G. H. *Trends Biochem. Sci.* **8**, 354–358 (1990).
13. Ward, W. H. J. & Fersht, A. R. *Biochemistry* **27**, 1041–1049 (1988).
14. Unwin, P. N. & Ennis, P. D. *Nature* **307**, 609–613 (1984).

The Arctic as a bellwether

John E. Walsh

A SMALL but significant decrease in the extent of Arctic sea-ice cover between the years 1978 and 1987 is reported on page 33 of this issue¹ by Gloersen and Campbell. The northward recession of the ice margin, revealed by satellite observations at two-day intervals, was accompanied by a reduction in the area of open water within the ice boundary. The area poleward of the ice boundary decreased by 2.1 per cent. By contrast, sea-ice coverage in the Antarctic showed no significant trends.

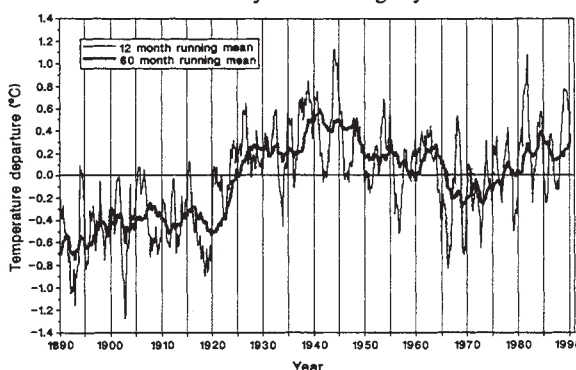
Because changes in the extent of snow and ice have long been regarded as manifestations and/or agents of climate change, the apparent decrease of Arctic ice extent must be viewed with interest. Indeed, global climate models run with doubled concentrations of CO₂ predict that the strongest greenhouse warming will occur in the polar regions. The polar amplification of the warming in model simulations has raised the possibility that greenhouse effects may become apparent sooner in the Arctic than elsewhere in the Northern Hemisphere. But such a picture of global warming is far from certain. Nevertheless, the recent literature does provide some ammunition for its proponents.

Studies reported during the past year have indicated that sea ice northeast of Greenland was substantially thinner in 1987 than in 1976 (ref. 2) and that the snow-covered area of the Northern Hemisphere's continents was smaller during 1990 than at any other time since homogeneous snow data became available in 1972 (ref. 3). There is also evidence⁴ that the springtime disappearance of snow cover over Alaska occurred about two weeks earlier in the 1980s than in the 1940s and 1950s. Data from lakes in central Canada⁵ show a warming of several degrees Celsius and a lengthening of the ice-free season by several weeks from the late 1960s to the late 1980s. Finally, the hemispheric asymmetry of the decrease of sea-ice extent found by Gloersen and Campbell is consistent with that of the greenhouse warming obtained in a recent climate model simulation by Stouffer, Manabe and Bryan⁶.

However, there are some important caveats and counter-indications that must be addressed before one concludes that the Arctic is heralding a greenhouse warming. First, the recent Arctic warming is neither an amplification of a global warming nor has it occurred throughout the northern high latitudes. The 'consensus' estimate of the increase of the global mean surface temperature during the past century is about 0.5 °C, but the temporal increase is not monotonic. Northern Hemisphere land-station temperatures show a warming from approximately 1900 to the 1940s, a cooling from the 1940s through the mid-1960s, and a warming

thereafter; Arctic temperatures have varied similarly (see figure).

The warming of the early 1900s was considerably stronger in the Arctic than elsewhere, but the warming since the mid-1960s is comparable in the Arctic and the mid-latitudes⁷. Although the recent Arctic warming is strongest over Alaska and northern Asia, some high-latitude areas such as northern Europe, the Labrador Sea and southern Greenland have actually cooled slightly^{7,8}.



Areally averaged temperature anomalies for the region poleward of 55° N. Thin line is 12-month running mean, thick line is 60-month running mean. (Land data from P. D. Jones, University of East Anglia; sea-surface temperature data from UK Meteorological Office.)

since the 1960s. (It is nonetheless worth noting that the more recent climate model simulations⁹ employing interactive oceans have substantially reduced the projected warming over the North Atlantic and northwestern Europe.) The relative warming of Alaska and northwestern Canada during the past 10–15 years is at least partially attributable to an increase of warm advection resulting from changes in the circulation of the lower atmosphere¹⁰, although there are indications¹¹ that the coldest airmasses affecting these regions are now several degrees warmer than they were 30–40 years ago.

Some of the Arctic changes mentioned earlier may also be manifestations of interannual variability or low-frequency atmospheric fluctuations unrelated to the greenhouse effect. For example, the differences of ice thickness between 1976 and 1987 noted by Wadhams² have been reproduced by a sea-ice model driven by observational analyses of interannually varying wind and temperature; but the differences obtained from the same model vanish if the initial (1976) or final (1987) dates are shifted by one or two years¹². Moreover, the occasional El Niño occurrences in the equatorial Pacific Ocean can cause significant anomalies of the atmospheric circulation and perhaps of sea ice over timescales of several years. As there were two such events during the 1978–87 period studied by Gloersen and Campbell, one cannot dismiss the possibility that the trend of the nine-year time series is attributable to the

phasing of the El Niño effect on the large-scale atmospheric circulation.

On slightly longer timescales, the warming of the North Pacific subarctic during the 1970s and 1980s has been attributed at least partially to very low frequency (20–30 year) fluctuations that are possibly associated with lunar tides or solar activity¹³. If such effects extend to sea-ice coverage, an apparent but temporary trend can appear in a time series spanning only a decade.

Finally, there are several reasons why the polar amplification of the CO₂-induced warming projected by climate models of the 1980s may have been exaggerated. Ice-albedo-temperature feedback contributes to

the polar amplification of the simulated warmings. The models that project polar warmings of 10–15 °C during winter (with doubled CO₂) generally treat sea ice as a uniform slab rather than as a mixture of ice and open water. The ocean-to-atmosphere heat flux in ice-covered areas during winter is underestimated by these models. The CO₂-induced warming and the resulting ice melt therefore create an artificially large enhancement of the heating of the atmospheric surface layers in high latitudes. This 'thermal shock' is further exaggerated in models that overestimate the

ice extent of the present climate.

It is also likely that the high-latitude cloudiness will increase as frozen sea ice is replaced by open ocean or melting ice. This increase of cloudiness may mask the surface albedo change and hence reduce the albedo-temperature feedback (climate model simulations of high-latitude cloudiness have generally been poor). A final consideration is that ice melt and/or increased precipitation in high latitudes are likely to freshen the surface waters of the subpolar seas, thereby reducing the 'thermohaline' convection that now brings heat from the deeper waters to the surface and the lower atmosphere over the subpolar seas. This oceanic response to a surface freshening has already appeared in recent climate model simulations⁹. The weakening of the thermohaline circulation results in a cooling (or a reduction of the warming) of several degrees

- Gloersen, P. & Campbell, W. J. *Nature* **352**, 33–36 (1991).
- Wadhams, P. *Nature* **345**, 795–797 (1990).
- Robinson, D. A. & Dewey, K. F. *Geophys. Res. Lett.* **17**, 1557–1560 (1990).
- Foster, J. L. *Arc. Alpine Res.* **21**, 60–70 (1989).
- Schindler, D. W. et al. *Science* **250**, 967–970 (1990).
- Stouffer, R. J., Manabe, S. & Bryan, K. *Nature* **342**, 660–682 (1989).
- Hansen, S. & Lebedeff, S. *J. geophys. Res.* **92**, 13345–13372 (1987).
- Jones, P. D. *J. Clim.* **1**, 654–660 (1988).
- Manabe, S. & Stouffer, R. J. *J. Clim.* **1**, 841–866 (1988).
- Trenberth, K. E. *Bull. Am. Meteor. Soc.* **71**, 988–993 (1990).
- Kalkstein, L. S. et al. *J. Clim.* **3**, 1153–1167 (1990).
- Walsh, J. E. & Chapman, W. L. *Proc. Fifth AMS Conf. Clim. Var.* (in the press).
- Royer, T. C. *J. geophys. Res.* **94**, 18,175–18,183 (1989).

in the subpolar North Atlantic. The warming of the circumpolar ocean around Antarctica is delayed by strong vertical mixing in a simulation with a similar model⁶.

These caveats suggest that the recent variations detected in the Arctic may well not be manifestations of a greenhouse-induced warming. Nevertheless, the fluctuations are generally consistent among the different

variables of the Arctic climate system, and they are qualitatively consistent with the changes that are expected to result from the increasing concentrations of greenhouse gases. Clearly, the Arctic will bear close watching over the next decade or two. □

John E. Walsh is in the Department of Atmospheric Sciences, University of Illinois, Urbana, Illinois 61801, USA.

DNA SEQUENCING

Complementary questions

Peter Little

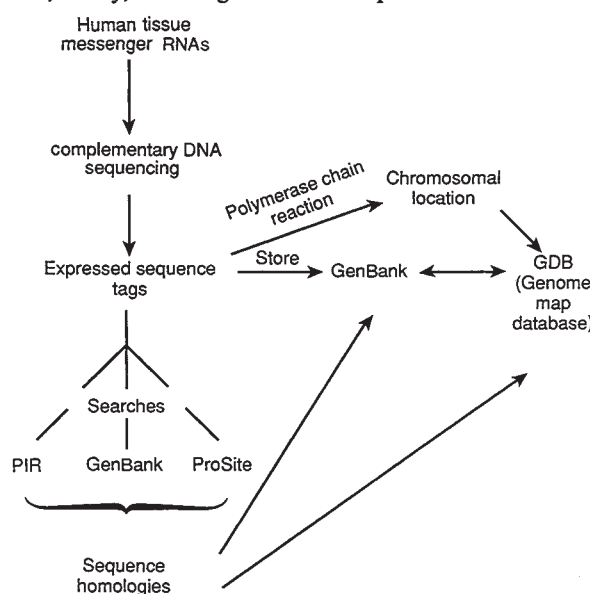
THE human genome project, with its ultimate goal of establishing the base sequence of all 3×10^9 base pairs of human DNA, has evolved in response to much critical comment, not least over the strategies to be employed. Writing in *Science*¹, Adams *et al.* report the first published results from a new approach to the central problem of DNA sequencing within the project. Underlying this work, however, are growing differences of opinion over whether the storage of DNA sequences is the only goal of the project, or whether the distribution to the scientific community of actual clones as reference libraries should also be one of the main activities.

An early objection to 'total genome' sequencing was that as a great deal of DNA does not code for proteins or structural RNAs, much of the sequence so obtained would be 'junk'. The 'complementary DNA strategy' — to sequence only transcribed DNA — was designed simultaneously to answer critics and generate biologically relevant DNA sequences at a higher rate and more cheaply than whole genome approaches. The strategy was developed independently by groups in the United States, particularly under J. C. Venter (at the National Institutes of Health in Bethesda), and in Britain under S. Brenner (Medical Research Council, Cambridge, and the UK Human Genome Resource Centre).

Venter's group¹ now describe the generation of sequences from brain cDNA libraries (it is logical to use brain tissue because 30–60 per cent of all messenger RNAs are expressed in this organ). They have sequenced 609 cDNAs, either by random sampling of cDNA libraries derived from the hippocampus, fetal brain or the temporal cortex, or by identifying potential brain-specific transcripts by selecting against those cDNAs that also appear in lung libraries. Automatic DNA sequencers were employed to generate 300–400 bases of sequence, which they call ESTs (expressed sequence tags) in distinction to the random and more general STSs (sequence tag sites)².

Expressed sequence tags are used to screen three sequence databases (GenBank, containing virtually all published DNA sequences from any species; PIR, containing protein sequences from all origins; and Pro-

Site, which is a specialist 'motif' database). These analyses identify five classes of DNA homologies within the ESTs — mitochondrial genes, repetitive DNAs, ribosomal sequences, normal human nuclear genes and, finally, nuclear genes of other species.



The complementary DNA strategy, and its relationship to existing databases.

Thirty six cDNAs had over 97 per cent sequence similarity with published human gene sequences and 54 had motif homologies. Crucially, 230 (38 per cent) had no matches and must be previously unknown genes; 134 cDNAs generated no data for technical reasons.

A great deal of effort in human genetics is being directed to analysis of phenotype by genetic mapping, and it is essential that cDNA sequences should be incorporated into the genetic map because it will make searching for candidate gene sequences a lot easier. Adams *et al.* recognize this and, using the polymerase chain reaction, make primer pairs off the ESTs: these are used to amplify DNA from a panel of somatic cell hybrids that contain different human chromosomes. The pattern of amplification gives the chromosomal location at low resolution.

These experiments are the start of a systematic project, displayed in the figure, to sequence expressed genes in humans. The

sequences are available as part of the DNA, protein and genetic databases routinely used by all molecular biologists, and the very reasonable hope is that by accumulating human cDNA sequences the 'hit' rate of matching independently isolated DNAs to database sequences will increase. If it does, it will have been done more efficiently, rapidly and cheaply than the total genome approach.

This is a considerable body of work and one that is expanding rapidly. Venter's group (C. Venter, personal communication) has sequenced a further 1,000 human clones and 500 each from *Caenorhabditis elegans* and *Drosophila melanogaster*. Similarly, the British project has data on 1,200 sequences (R. Sibson, personal communication) from liver, placental, brain and rhabdomyosarcoma cDNA libraries. The results of the sequence analyses are very similar to those obtained by Venter.

Several technical obstacles remain, however. Many mRNAs are present in more than a single copy per cell and so much time could be wasted sequencing different cDNAs derived from the same gene (though the construction of 'normalized' libraries, with each transcript represented once in each library, is under investigation both in Britain and the United States). It will also be difficult to avoid some redundancy of effort in sequencing all the different libraries that will be needed to approach complete coverage of all expressed genes.

Expressed sequence tags stored in a database, with chromosomal region defined, are not all that is required of the genome project. In Europe (and increasingly the United States) there are systematic attempts to distribute DNA

clones to any laboratory that requests them³. The DNA samples of the clones are immobilized on filters in a systematic array which can be hybridized with any probe that the receiving laboratory wishes. These arrayed clones are called reference libraries. Because all laboratories get the same array, it is simple for the distributing agent to record the results of these experiments and draw cross relationships (for example, one clone hybridizing with a motif probe in one lab and another cDNA in a second) that would not have been available to either group independently⁴.

There are four principal features of the reference library strategy: hybridization screens all the clones for cross reaction, not just those that have been sequenced; it generates information that is additional to partial DNA sequences; it gives researchers an immediate source of material; and, finally, it places the onus of work on the laboratory most interested in the results — the specific end user. If a particular laboratory deems a particular

experiment essential, it will put resources and expertise into clone isolation that would be meaningless if applied to the same cDNA in a random context.

In their paper, Adams *et al.* state that "Sequence tagged sites are becoming standard markers for the physical mapping of the human genome". This is certainly a widely held view, but STSs are a standard and not the standard. As a practical demonstration of reference library use, Lehrach's group in London has for several years distributed chromosome-specific cosmid libraries from chromosomes X and 21, and latterly 17 and 22 (refs 3 and 4). They have passed out 365 clones to 83 laboratories (H. Lehrach, personal communication). This is a clear demonstration of demand. At present the libraries are of genomic clones, and will soon include a total YAC library, but there are no conceptual differences from cDNA library distribution.

HIGH-ENERGY PHYSICS

Evidence of supersymmetry

G. G. Ross

NATURE may respect a new, hitherto unseen symmetry, called supersymmetry, and there may be a new form of matter, according to several new papers¹⁻³. The evidence advanced is indirect and relies both on the theoretically motivated idea that the fundamental forces of nature are all aspects of a 'grand unified theory' (GUT) and on the laboratory measurements of the strengths of these forces, including the recent high-precision measurements at CERN's Large Electron Positron Collider of the properties of the Z^0 boson, the mediator of the neutral weak interaction. The implications of such a result are dramatic for, if true, we are on the threshold of discovery of a rich new world of supersymmetric particles and closer to the ultimate 'theory of everything'.

The standard model of the strong, electromagnetic and weak interactions suggests all three forces have a common origin, mediated by spin-one (vector) bosons coupled to matter in the form of spin-one-half fermions, the quarks and leptons. Thus the electromagnetic interaction is mediated by the photon coupled to electric charge, the strong interactions are mediated by eight vector gluons coupled to a new colour charge carried by the quarks, and the weak interactions are mediated by three vector bosons, the $W^\pm$ and Z^0 bosons, coupled to weak charge carried by the quarks and leptons. However, the standard, unified model does not explain the origin of the different types of charge nor the fact that the couplings associated with the different interactions are quite different. Grand unification was invented to solve these problems and does so by relating the vector bosons and charges by an enlarged symmetry which has a single fundamental coupling. The archetype GUT is based on

Pressure is coming onto the British cDNA strategy to distribute filters, and it is to be hoped that the same will happen in the United States. The idea that STSs will be the only sort of stored, cloned DNA resource is not the final solution. In my view, the success or failure of the reference library compared to STS approaches will be defined in the most simple of ways — the end user scientist will dictate. The funding bodies can either respond or explain why the customer is, for once, wrong. □

Peter Little is in the Department of Biochemistry, Imperial College, Prince Consort Road, London SW7 2AZ, UK.

1. Adams, M. D. *et al.* *Science* **252**, 1651-1656 (1991).
2. Olson, M., Hood, L., Cantor, C. & Botstein, D. *Science* **245**, 1434-1435 (1989).
3. Nizetić, D. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **88**, 3233-3237 (1991).
4. Lehrach, H. *et al.* in *Genome Analysis*, Vol. 1 *Genetics and Physical Analysis* (eds Davies, K. E. & Tilghman, S.) 39-81 (Cold Spring Harbor Laboratory Press, 1991).

the mathematical group $SU(5)$ which relates the five fundamental charges needed to describe the colour, electric and weak charges. This relation translates to a relation between the quarks and leptons, these being grouped together in simple representations of $SU(5)$.

Initially this elegant idea foundered on the fact that the (suitably normalized) couplings of the fundamental interactions are predicted in $SU(5)$ to be the same, grossly in

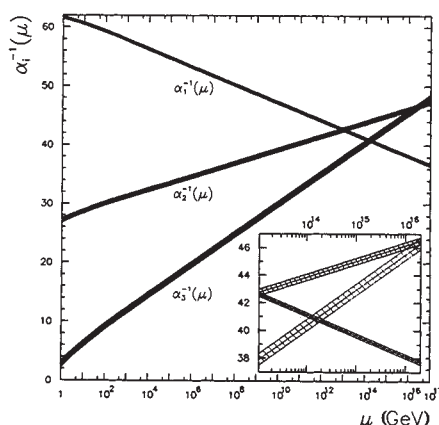


FIG. 1 Evolution with respect to the measurement scale, μ , of the strong coupling, α_s , and the electroweak couplings α_2 , α_1 , associated with the $SU(3) \otimes SU(2) \otimes U(1)$ gauge group of the standard model (from Amaldi *et al.*¹). The effect of the radiative corrections is to make α_i evolve almost linearly in $\log(\mu/M_x)$ as the measurement scale μ is varied. The couplings come together at high energies in qualitative agreement with the predictions of grand unification that they should all be equal at the GUT scale M_x . But the prediction fails quantitatively, thus ruling out simple GUTs such as $SU(5)$, $SO(10)$ or E_6 which break in a single step to the standard model.

contradiction with experiment. The situation is drastically changed, however, by the inclusion of corrections to the couplings coming from the self interactions of the vector bosons and their interaction with the quarks and leptons⁴. Because the number of interacting states is different for the strong, weak and electromagnetic interactions these 'radiative' corrections cause their couplings to differ, the difference increasing as the energy or mass scale at which they are measured, μ , drops further below the scale M_x at which $SU(5)$ is broken. Below this scale the symmetry between the five charges is lost, leaving residual symmetries relating the colour, weak and electric charges separately (via the group $SU(3) \otimes SU(2) \otimes U(1)$).

This is illustrated in Fig. 1 where it can be seen that the qualitative behaviour is in the correct direction; the strong coupling is larger than the other couplings at the low energy scales at which the laboratory measurements of the couplings are made. But the radiative corrections depend only logarithmically on the ratio μ/M_x , and to bring the predictions in line with experiment, a very large value for M_x is needed — of the order of 10^{15} gigaelectronvolts (GeV), some fifteen orders of magnitude greater than the proton mass. Such a large scale of grand unification proved to be fortuitous for it was found that, owing to the unification of quarks and leptons, GUTs typically violate baryon- and lepton-number (they can be destroyed) at a rate inversely proportional to the fourth power of M_x . Proton decay has never been observed and the present experimental limits do require a very large scale for M_x .

Although this qualitative picture is encouraging, the detailed quantitative comparison with $SU(5)$ fails. Using the new precision measurements for the weak couplings from the Large Electron-Positron Collider (LEP) and from neutrino scattering experiments together with the accurately known electromagnetic coupling and the somewhat less accurately measured value for the strong coupling, several authors¹⁻³ have now shown that the radiatively corrected couplings fail by seven standard deviations to cross at a single high-energy point (Fig. 1). Moreover the best value for M_x , although very large, still leads to an unacceptably fast rate of proton decay.

The suggested resolution to this problem lies in a modification of the GUT to include a new symmetry. This symmetry, known as supersymmetry, relates integer-spin bosonic states to half-integer spin 'fermionic' states (see ref. 5 for a review). Supersymmetry has long been considered as a possible symmetry of nature and appears to be essential if the unification of the fundamental forces is to include gravity. Moreover, studies of GUTs have suggested they are inconsistent without supersymmetry because, in non-supersymmetric theories, radiative corrections to masses drive the electroweak breaking scale (at which weak and electromagnetic forces

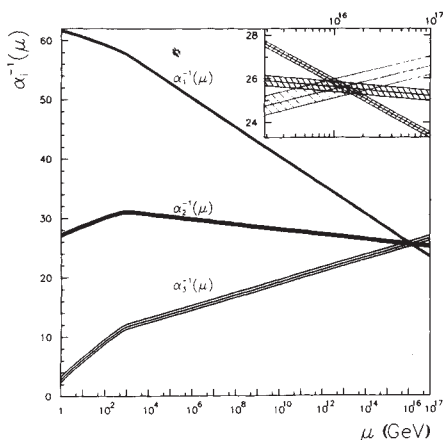


FIG. 2 The evolution of the couplings in the supersymmetric version of the standard model. The fact that the couplings meet at a point is in agreement with simple GUT predictions such as supersymmetric SU(5) which break in a single step to the standard model; in this case M_X is large enough adequately to inhibit proton decay⁹. The break in the slope of the curves corresponds to the threshold for the new supersymmetric states and agreement with the GUT predictions is maintained only if the masses of the new supersymmetric states lie in the range 10^2 – 10^4 GeV.

become distinct) and the associated $W^\pm$ and Z^0 masses up to M_X , some thirteen orders of magnitude too large.

This implies that a self-consistent theory of Grand Unification requires an extension of the standard model to include supersymmetry to protect it against such radiative corrections. Moreover to be effective, the symmetry must not be broken at an energy scale much larger than the electroweak breaking scale and hence, as I will discuss next, new supersymmetric states should be relatively light and accessible to experimental detection.

The extension of the standard model to include supersymmetry requires the addition of many new states. These include the 'squarks' and 'sleptons', spin-zero partners of the spin-one-half quarks and leptons needed in the standard model. Also needed are the gluinos, the Wino, the Zino and the photino, spin-one-half partners of the gauge bosons of the standard model, the gluon, the $W^\pm$, the Z^0 and the photon. Although these new states are thought to be considerably heavier than their partners, and hence should not yet have been found directly in laboratory experiments, they will contribute to the radiative corrections. The exciting observation is that their inclusion brings the predictions of grand unification into remarkable agreement with experiment.

The effect of supersymmetry on grand-unified predictions has been known for many years^{6–8}. Initially the non-supersymmetric SU(5) predictions seemed to be favoured but the recent experimental improvement in the precision measurements of the gauge couplings, largely through the precision measurements of the Z^0 -boson coupling at LEP, has led to the conclusion that the supersymme-

tric prediction is quantitatively better than the non-supersymmetric result. This may be seen in Fig. 2 where, in addition, it may be seen that the precision is now so good that we may even estimate the mass that the new supersymmetric states must have if their radiative effects are to bring the prediction into agreement with experiment. Amaldi *et al.*² find that the squarks, sleptons and gauginos should have an average mass in the range 10^2 – 10^4 GeV. Thus one may argue that there is evidence for new forms of (relatively) light supersymmetric matter and an underlying grand unified theory.

The result is so dramatic that it needs to be evaluated with some caution. The most obvious reservation is that it relies on the theoretical extrapolation of the standard model, albeit with the inclusion of supersymmetry, twelve orders of magnitude beyond the energy scale at which it has been tested. No theory has proved to be so robust in the past, so it is understandable if one views this extrapolation with some caution. Even given the framework of grand unification there is considerable uncertainty, for the relative values of the couplings may differ in different unification schemes. Within the (large) class of theories that give the SU(5) predictions, there are corrections coming from virtual states with mass, around or greater than M_X which are not included in the analysis and which can affect the results substantially. Nonetheless it is remarkable that the simplest possible extension of the standard model to include supersymmetry coupled with the simplest assumption about grand unification yield predictions in

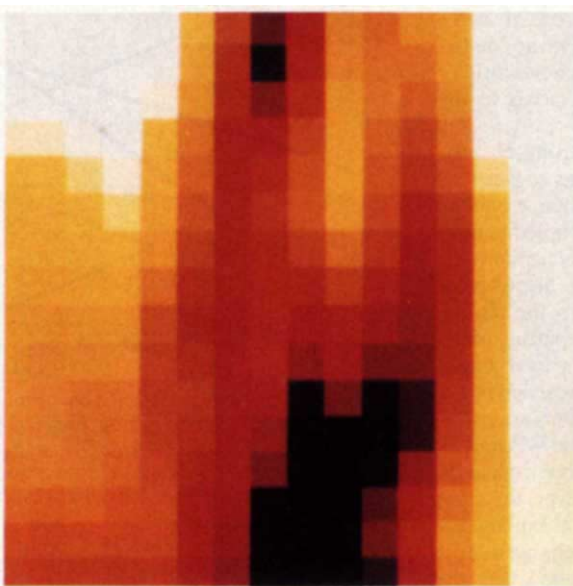
detailed agreement with experiment and require a very low scale of mass for the new supersymmetric states.

What are the implications of the predicted spectrum of supersymmetric states? If the lower value of the allowed masses is correct, some supersymmetric states should be accessible to the present accelerators, LEP at CERN or Fermilab in Chicago. So far searches have not produced any evidence of such states. The favoured value of 10^3 GeV would put the supersymmetric states out of the reach of these machines but firmly in the range of future machines, the Superconducting Super Collider in Texas and the proposed Large Hadron Collider (LHC) at CERN. If so, the new accelerators will be able to open a window on the new supersymmetric world and, indirectly, on the ultimate 'theory of everything' that many physicists think will unite the fundamental forces including gravity. □

G. G. Ross is in the Department of Physics, University of Oxford, 1 Keble Road, Oxford OX1 3NP, UK.

1. Ellis, J., Kelley, S. & Nanopoulos, D. V. *Phys. Lett.* **B260**, 131–137 (1991); **B249**, 441–448 (1990).
2. Amaldi, U., deBoer, W. & Fürstenau, H. *Phys. Lett.* **B260**, 447–455 (1991).
3. Langacker, P. & Luo, M. Univ. Pennsylvania Preprint UPR-0466T (1991).
4. Georgi, H., Quinn, H. R. & Weinberg, S. *Phys. Rev. Lett.* **33**, 451–454 (1974).
5. Ross, G. G. *Grand Unified Theories* (Benjamin/Cummings, 1984).
6. Dimopoulos, S. & Georgi, H. *Nucl. Phys.* **B193**, 160 (1981).
7. Ibanez, L. E. & Ross, G. G. *Phys. Lett.* **B105**, 439 (1982).
8. Einhorn, M. & Jones, D. R. T. *Nucl. Phys.* **B196**, 475–488 (1982).
9. Dimopoulos, S., Raby, S. & Wilczek, F. *Phys. Rev.* **D24**, 1681–1683 (1981).

Nobody nose . . .



THIS is what ethanol looks like to the simple 'electronic nose' described by Lundström *et al.* on page 47. Other gases, such as ammonia and hydrogen, paint a different cubist picture, so that each leaves a distinctive fingerprint. The image is a representation of the variation, caused by the absorption of ethanol, in surface potential of a sensing surface comprising adjacent strips of platinum, iridium and palladium. Top-to-bottom variations, resulting from an imposed temperature gradient, add to the information in the image. The authors make the physiological analogy on the basis of one of the prevailing views on mammalian olfaction: that the

10,000 or so distinguishable odours are identified by a much smaller number of receptors which send signals to specific locations in the brain, producing a type of spatially resolved neuronal fingerprint. In the device of Lundström *et al.*, some decoding of the image is possible in terms of known chemical behaviour: the low affinity of iridium (the central strip) for ethanol, for example, is clearly visible. □

Designing a slow leak

Richard B. Kaner

IF LARGE surface areas could be coated efficiently with pinhole-free polymer films of less than 50 nm in thickness, many important membrane-based technologies such as gas separators and sensors would be greatly improved. Langmuir–Blodgett and interfacial polymerization techniques have so far failed to produce nonporous ultrathin films on porous substrates, although ultrathin film coatings on smooth substrates are now possible¹. C. Liu and C. R. Martin report on page 50 of this issue² an important advance in coating porous substrates by describing a method for growing ultrathin coherent films using photoinitiated polymerization. Capillary action is used to cover the surface of a porous material with a thin film of monomer solution. Irradiation with ultraviolet light then produces the high-quality films.

Scanning electron micrographs of the thinnest polymer films reported by Liu and Martin indicate uniform 40-nm thicknesses. The authors measured the gas permeabilities of oxygen and nitrogen to determine the film's porosity: pinholes would allow the gases to flow by Knudsen diffusion (at a rate inversely proportional to the square root of their masses) and the lighter element, nitrogen, would permeate 1.07 times faster than oxygen. But, oxygen actually permeated up to eight times faster than nitrogen, so that the films were defect-free, and can be used for gas separation.

Gas-separation techniques have evolved over the past 15 years into a large commercial enterprise, involving at least 20 major companies worldwide. The driving forces behind replacing conventional cryogenic gas separation by membrane-based methods are the potential reduction in energy consumption and capital costs and the greater ease of operation at remote plant sites. The largest current market is for nitrogen separated from air, for use in protecting perishable goods and blanketing inflammable materials. The oxygen-enriched byproduct can be used to accelerate combustion or for medical purposes. Other applications include recovering and recycling hydrogen during ammonia synthesis and petroleum refining. Carbon dioxide separated from natural gas can be reinjected into wells to assist petroleum recovery. Future applications could include pollution control.

For effective gas separation, membranes must both be highly selective and allow high flux rates. Flux is inversely proportional to the thickness of the membrane, so that the thinnest possible nonporous polymer is desirable. But the polymer must also withstand high pressures (up to 150 atmospheres), and separation is generally carried out with asymmetric membranes in which a thin skin of polymer is grown on a porous structural support. In 1979, workers at

Monsanto Company introduced self-supporting hollow-fibre membranes³. These asymmetric tubules had a porous polymer support to withstand high pressures, a thin polymer skin to carry out separation and large surface areas to allow high gas throughput. The hollow fibres had walls 25–250 µm thick with a dense skin of 0.1–1 µm. Clearly if the separating layer could be reduced to 40 nm, as Liu and Martin have achieved, much higher flux rates could be possible. The real difficulty, yet to be tackled, will be producing ultrathin films over enormous areas without introducing pinholes. A typical gas-separation module (6 feet by 10 inches diameter) contains thousands of tubules with thousands of square feet of surface area.

Ultrathin films could also be used in sensors⁴. For example, surface-acoustic-wave devices can detect mass changes of less than one nanogram per square centimetre, by exploiting resonant-frequency changes in piezoelectric crystals. Sensitive detectors with rapid responses could be made by adding a coating film of selectively absorbing polymer. In this type of sensor, the thickness is the main limitation to response time. Other applications of ultrathin films may include bioreactors or drug-delivery systems⁵.

Liu and Martin indicated that they can synthesize ultrathin films of polymers with the ability to exchange ions, or that have photoactive or electroactive properties. The

latter types include conjugated, conducting polymers, which have been investigated over the past 15 years for use as active components in batteries, electrochromic displays and solar cells, or as antistatic coatings, electromagnetic shields and 'smart' windows⁶. Even gas sensors which operate at room temperature and can detect minute quantities of hydrogen sulphide or ammonia in air are possible⁷. Recently we reported⁸ that the doping process that makes polymers conducting can be used to induce permanent morphological changes in free-standing, environmentally stable polyaniline films. This process results in even higher gas selectivities for oxygen/nitrogen, hydrogen/nitrogen and carbon dioxide/methane. If ultrathin film concepts can be combined with the superior gas selectivity of conducting polymers many novel devices should be forthcoming. □

Richard B. Kaner is in the Department of Chemistry and Biochemistry, University of California, Los Angeles, California 90024, USA.

1. Gouchanour, C. 4th natn. Meeting North Am. Membrane Soc. San Diego, California, May 29–31 (1991).
2. Liu, C. & Martin, C. R. *Nature* **352**, 50–52 (1991).
3. Henis, J. M. S. & Tripodi, M. K. *Science* **220**, 11–17 (1983).
4. Ward, M. D. & Buttry, D. A. *Science* **249**, 1000–1007 (1990).
5. Tierney, M. J. & Martin, C. R. *J. electrochem. Soc.* **137**, 2005–2006 (1990).
6. Kaner, R. B. & MacDiarmid, A. G. *Scient. Am.* **258**, 106–111 (1988).
7. Miasik, J. J., Heeger, A., Moseley, P. T. & Tofield, B. C. in *Conducting Polymers Special Applications* (ed. Alcácer, L.) 189–198 (Reidel, Dordrecht, Holland, 1987).
8. Anderson, M. R., Mattes, B. R., Reiss, H. & Kaner, R. B. *Science* **249**, 1412–1415 (1991).

CYSTIC FIBROSIS

Acidification indication

Michael J. Welsh

Cystic fibrosis is caused by mutations in a single gene encoding the cystic fibrosis transmembrane conductance regulator (CFTR)¹. Although recent studies suggest that CFTR is a Cl[−] channel^{2,3}, it is not yet clear how this accounts for the numerous abnormalities of the cystic fibrosis phenotype. On page 70 of this issue⁴, Barasch *et al.* present a hypothesis to explain how some of the abnormalities might arise.

Ten years ago, the literature on cystic fibrosis was in confusion, with descriptions of a huge variety of clinical manifestations⁵ and many bewildering research observations. It was difficult to develop a unifying hypothesis about the underlying defect, other than to note that the disease primarily afflicted epithelia. Interest began to focus on Cl[−] channels when Quinton⁶ discovered that the sweat-gland-duct epithelia of cystic fibrosis patients had decreased Cl[−] permeability and Knowles *et al.*⁷ made similar observations using airway epithelia. Subsequently, defective cAMP-dependent regulation of epithelial Cl[−] channels drew

increasing attention; after all, this defect was responsible for an all-or-none difference between normal and cystic fibrosis cells in several organs⁸. But more quantitative and tissue-specific observations remained enigmatic, such as increased Na⁺ absorption in airway epithelia and increased sulphation and reduced sialylation of glycoproteins.

With the discovery of the gene responsible for cystic fibrosis¹, and the realization that it encodes CFTR, work centred around how mutations in it might eliminate the function of cAMP-regulated Cl[−] channels. It seems that CFTR is itself a cAMP-regulated Cl[−] channel^{2,3}; if so, a loss of Cl[−] channel function is tied to the defective gene. This link begins to explain the pathophysiology in the sweat gland duct (a loss of Cl[−] channel function causes the reduced Cl[−] permeability and hence the failure to reabsorb Cl[−] from the sweat), the pathogenesis of meconium ileus (defective Cl[−] channel function would impair intestinal fluid secretion, thereby generating a dehydrated meconium), and the pathogenesis of pancreatic disease

(defective secretion in the ducts would give rise to a dehydrated pancreatic fluid that obstructs the ducts, thereby causing pancreatic atrophy). Our understanding of these problems, though, is far from complete.

But what about the pulmonary airways? They are the site of repeated infections in cystic fibrosis and are the main source of morbidity and mortality. Can the pathogenesis and pathophysiology be attributed entirely to a failure of transepithelial Cl^- secretion by airway epithelia? And what about the many different research observations that are not readily accounted for by CFTR being a Cl^- channel (and that have been used to argue that CFTR is not a Cl^- channel)? These questions are not trivial: to understand how mutations in CFTR cause disease, we must first understand how such mutations produce so many different phenotypic manifestations.

That is where the work by Barasch *et al.*⁴ fits in. The authors present an intriguing hypothesis by which the loss of Cl^- channel function could produce some of the phenotypic manifestations of cystic fibrosis. They propose that defective acidification of intracellular compartments, caused by defective Cl^- permeability, may be responsible for abnormalities of protein modification.

Certain intracellular vesicles are acidified by a proton pump and a Cl^- channel, located in parallel on the vesicular membrane. The proton pump provides the driving force for acidification and the Cl^- channel provides both a pathway for anion movement (to maintain electroneutrality) and a means to regulate acidification. It has been postulated that the proton pump is constitutively active and that the opening and closing of Cl^- channels regulates acidification. If the Cl^- channel does not open, then acidification becomes impaired. The results of Barasch *et al.* are the first to show that acidification of the *trans*-Golgi, endosomes and prelysosomes is defective in cystic fibrosis airway epithelia. The implication is that a defective vacuolar Cl^- channel is responsible.

The activity of *trans*-Golgi enzymes, such as sialyltransferases, is sensitive to changes in pH. Barasch *et al.* therefore compared sialylation of glycoproteins and glycolipids in transformed cells from either a cystic fibrosis nasal polyp or a normal trachea. They found that sialylation was reduced in the cystic fibrosis polyp cells. This is perhaps the most enticing observation — if defective acidification and sialylation can be demonstrated in several primary cultures and convincingly linked to abnormal Cl^- channel function, then we are well on our way to having established a defect in Cl^- channels that can induce changes in vesicle enzyme activity and thereby alter the properties of proteins and lipids. The mechanism would also provide a way in which a qualitative, all-or-none defect in Cl^- channels can lead to a quantitative defect, such as the increased protein sulphation seen in cystic fibrosis. And could it be that such altered processing of surface

proteins contributes to the colonization and adherence of microorganisms to the airway epithelium? In the end, infection of the airways is one of the main causes of morbidity.

Although the hypothesis is intriguing, there are many questions that remain unanswered. Is CFTR directly responsible for normal and abnormal acidification? If the answer is yes, then CFTR should be localized to both the affected intracellular membranes and to the plasma membrane where it causes defective apical Cl^- permeability. And as the Cl^- channel that is defective in cystic fibrosis is regulated by cAMP, does this compound have an effect on vacuolar acidification? A cAMP-regulated Cl^- conductance has been demonstrated in endocytotic vesicles from renal epithelia⁹; is that response due to CFTR or to a different cAMP-regulated Cl^- channel? (The kidney does not have much CFTR.) Is the response in kidney vesicles the same as that studied by Barasch *et al.*, and is it defective in cystic fibrosis? Every cell requires acidification of some intracellular vesicles. If acidification is defective in cystic fibrosis epithelia, one would presume that it is not defective in other organs, because the process is so critical to a variety of cell functions. Perhaps there are different mechanisms of acidification in nonepithelial cells and cells that do not express CFTR.

Another problem is explaining how a defect in CFTR induces an increased rate of Na^+ absorption in cystic fibrosis airway epithelia¹⁰. Increased Na^+ absorption might alter the respiratory tract fluid; hence therapeutic attempts have been made with inhaled amiloride to inhibit Na^+ absorption in cystic fibrosis patients¹¹. The molecular and cellular events that couple Cl^- channel defects to increased activity of Na^+ channels are completely unknown. One wonders whether CFTR, as well as possibly being a cAMP-regulated Cl^- channel, might have an additional, as yet undiscovered function.

There is much to learn about CFTR and how mutations in it cause disease. And cystic fibrosis has never given up its secrets easily. But as the mysteries unfold, we are certain to learn exciting new biology. More importantly, we trust that the new insights will lead to better approaches to therapy. □

Michael J. Welsh is in the Howard Hughes Medical Institute, University of Iowa College of Medicine, Iowa City, Iowa 52242, USA.

Flying doctors

MANY terrible diseases are transmitted by blood-sucking insects. Some of them, such as yellow fever, dengue, plague and certain forms of encephalitis, are caused by a virus. They cannot be treated, but can be prevented by prior vaccination. Daedalus now wants the insects to transmit, not the disease, but the vaccination.

Vaccination often uses a mutated 'attenuated' virus with mild and transient effects, or the virus of a related and harmless disease (like the cowpox used by Jenner in his pioneering vaccinations against smallpox). These days, genetically engineered viruses are also used. In all cases, the vaccination stimulates the production of antibodies which inactivate the real virus when it comes along.

So DREADCO biologists are loading mosquitoes, fleas and so on with vaccination viruses. Like the fake cows used to attract tsetse flies, their 'vaccination-zombie' looks and smells invitingly human, but has vaccine-loaded fake blood under its skin. The deluded insects bite this, and take up the vaccine.

Much biological cunning will be needed to ensure that the vaccine-virus survives in the insect, and can be transmitted reliably to its subsequent victims. But when the technique has been perfected, vast numbers of vaccination-insects will be released in the disease ridden regions. As in the 'sterile male' technique of combating pests such as the screw-worm fly, the treated insects will briefly overwhelm the local ecology. Since they all carry the vaccine, while relatively few of the wild insects actually carry the disease, vaccination should outpace infection throughout the whole area. The most dispersed, doctor-shy, or needle-averse population will soon be thoroughly bitten and vaccinated. Even animal reservoirs of the disease will be reached and treated. Heroic expeditions through bush and jungle by paramedics and insect-spraying teams will no longer be needed. The insects that previously spread the disease will now spread the treatment.

With luck, the vaccine-virus will become stably established in the human population, or perhaps in some local wild animal. Indigenous bloodsucking insects will then spread booster vaccinations regularly through the population. The disease will be kept permanently in check.

This cunning technique may even work against viral diseases not carried by insects, such as rabies, poliomyelitis and hepatitis B. An insect-compatible virus, genetically engineered to carry some antigen of the chosen disease, should do the trick. Only one problem baffles Daedalus — how to give the patients a bureaucratically valid vaccination certificate?

David Jones

1. Riordan, J. R. *et al.* *Science* **245**, 1066–1073 (1989).
2. Anderson, M. P., Rich, D. R., Gregory, R. J., Smith, A. E. & Welsh, M. J. *Science* **251**, 679–682 (1991).
3. Kartner, N. *et al.* *Cell* **64**, 681–691 (1991).
4. Barasch, J. *et al.* *Nature* **352**, 70–73 (1991).
5. Boat, T. F., Welsh, M. J. & Beaudet, A. L. in *The Metabolic Basis of Inherited Disease* Vol. 6 (eds Scriver, C. R., Beaudet, A. L., Sly, W. S. & Valle, D.) 2649–2680 (McGraw-Hill, New York, 1989).
6. Quinton, P. M. *Nature* **301**, 421–422 (1983).
7. Knowles, M., Gatzky, J. & Boucher, R. J. *clin. Invest.* **71**, 1410–1417 (1983).
8. Quinton, P. M. *FASEB J.* **4**, 2709–2717 (1990).
9. Bae, H.-R. & Verkman, A. S. *Nature* **348**, 637–639 (1990).
10. Boucher, R. C., Stutts, M. J., Knowles, M. R., Cantley, L. & Gatzky, J. T. *J. clin. Invest.* **78**, 1245–1252 (1986).
11. Knowles, M. R. *et al.* *New Engl. J. Med.* **322**, 1189–1194 (1990).

Did Sirius change colour?

SIR — Schlosser and Bergmann¹ claim that Babylonian and early mediaeval authors saw Sirius as red, which if true would represent an anomaly in current evolutionary theories of white dwarfs. But van Gent² and I³ independently suggested that it might be a case of mistaken identity, and that, on evidence from Chinese sources, the brightest star in our sky has been white all along. The 'Rubeola' (which means reddish) of Gregory in the sixth century should really be Arcturus rather than Sirius, as justified by McCluskey⁴ and van Gent⁵; although Schlosser and Bergmann⁶ were unconvinced, and provided equally detailed counter-arguments. Also, Bruhweiler, Kondo and Sion⁷ speculated that Sirius B underwent a thermonuclear runaway event, to explain its purported redness in antiquity. Gry and Bonnet-Bidaud invoked⁸ the star's transit behind a cold interstellar cloud as the cause of a supposed increase and then decrease in the degree of reddening. I suggest that there is no ground for believing in any colour change, sudden or gradual.

Support for a gradual variation in Sirius's colour⁸ comes from a Chinese work of the first century BC. The text appears in an earlier part of the chapter that contains the same passage cited by me³ to affirm the whiteness of Sirius. Column *b* of the figure literally states that:

Wolf changes colour/Thieves and robbers abound

where Wolf is the Chinese name for Sirius. Read in context, however, it really means that if you perceive Sirius in a different colour (from white), then you can make the astrological prediction that follows.

In attempting any modern interpretation of ancient scientific writings, we must interpret their statements in the context of the dominant models of that time for that culture. In this case, the model is for inferring world events from astronomical observations, a consequence of the philosophical conception that the cosmos, the human society and individual human bodies respond to one another in turn. Thus, unlike Hellenism, Chinese tradition viewed the celestial sphere as full of secular variations in display, and indeed wanted it that way. A colour modification in a star could therefore be a subjective assessment — if you recited the right chant then you would recognize it — or some meteorological phenomenon. An interesting parallel is the discussion, attributed to an Egyptian source of 2 BC, of omens revealed by the apparent colour of Sirius at its heliacal rising².

To be convinced of the hypothetical nature of the colour change in the Chinese writing in question, we need look no further than its preceding passage (*a* in the figure), which says:

South of [Betelgeuse] are four stars called the Celestial Commode/Below the Com-

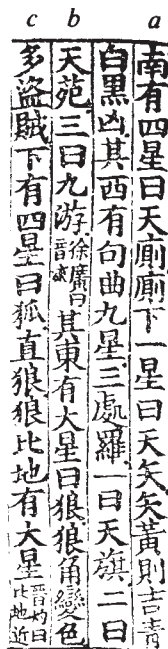
mode is one star call the Celestial Stool/Stool is yellow, everything fine/Green, white or black, it spells danger/To the west...

The Stool is the Orion Nebula, but surely any historical colour changes could only be imaginary.

One last cautionary note on text-based archaeoastronomical research: namely the danger of relying on a single excerpt from one extant work. It may have been incorrectly written, or errors introduced during repeated copying or reprinting. An example is provided by the passage that follows immediately (*c* in the figure):

Beneath are four stars called the Bow/[With an Arrow] pointing at the Wolf.

The number is wrong. The *History of Jing Dynasty* mentions nine stars instead, and indeed it would be stretching the analogy a



Excerpt from a Chinese book of the Western Han dynasty (Northern Sung edition). *a-c*, Passages referred to in the text.

bit too far to make a bow and arrow out of merely four points. I maintain³ that the fact that Sirius looked white was not contradicted in any of the other dynastic histories compiled at later times.

TONG B. TANG

Department of Physics,
Hong Kong Baptist College,
Kowloon, Hong Kong

- Schlosser, W. & Bergmann, W. *Nature* **318**, 45 (1985).
- van Gent, R. H. *Nature* **312**, 302 (1984).
- Tang, T. B. *Nature* **319**, 532 (1986).
- McCluskey, S. C. *Nature* **325**, 87 (1987).
- van Gent, R. H. *Nature* **325**, 87 (1987).
- Schlosser, W. & Bergmann, W. *Nature* **325**, 89 (1987).
- Bruhweiler, F. C., Kondo, Y. & Sion, E. M. *Nature* **324**, 235 (1986).
- Gry, C. & Bonnet-Bidaud, J. M. *Nature* **347**, 625 (1990).

Bacteria mating preferences

SIR — Although the parasexual processes of bacteria are very different from the sexual systems of higher eukaryotes, there are often surprising parallels. In particular, the preferential uptake of homologous transforming DNA by competent bacteria closely resembles the mating preferences exhibited by many animals (see ref. 1 for a recent review).

When naturally transformable bacteria become competent, they can efficiently take up large DNA fragments which can recombine with and replace homologous segments of the genome. Competent gram-negative bacteria preferentially take up DNA fragments containing specific 9–11-base-pair sequences (AAGTGC GGTCa in *Haemophilus influenzae*²; GCCGTCTGAA in *Neisseria gonorrhoeae*³). Because their genomes contain many dispersed repeats of the preferred sequences, DNA from the same species is taken up much more rapidly than foreign DNA.

The correspondence between uptake preference and sequence abundance must result from selection. There are about 100 times more copies of the uptake sequences than expected in the absence of selection, and their presence in a fragment increases its uptake at least several fold. Transformation thus provides an example of an evolved mating-preference system. In the framework of sexual mating systems, competent cells can be viewed as females discriminating between potential mating partners (DNA fragments), with noncompetent cells being the asexual precursors of both females and free DNA.

The evolution of DNA uptake specificity in transformable bacteria has not yet been subject to the theoretical and experimental analysis applied to sexual mating preferences. One simple model might assume that the DNA receptor initially had a minor unselected sequence bias. DNA fragments containing newly arisen copies of the preferred sequence would then be more efficiently taken up, and could physically replace the resident wild-type sequence by recombination. By its selection of fragments the receptor would thus exert a recombination pressure that gradually increased the frequency of the preferred sequence in the genome. The increased number of copies would then confer a selective advantage on cells whose receptors had a higher affinity for the sequence, and the increased receptor specificity would further increase the recombination of new copies into the genome, until the feedback was limited by other factors such as coding requirements.

The driving force in this model is the high affinity of sequence-specific DNA binding, a direct advantage to the competent cell. This benefit is independent of whether the primary function of transformation is sexual (production of genetically recombinant progeny) or physiological (provision of nucleo-

tides for nutrition or of templates for DNA repair), as long as most of the available DNA is from the same species and there is occasional recombination between incoming DNA and the chromosome. Indeed, if acquisition of nucleotides is the function of transformation then any DNA will serve, but the feedback between sequence abundance and receptor specificity will only occur if some of the incoming DNA recombines with the chromosome. The model has not, however, been rigorously examined, and is only one possible explanation for uptake specificity.

ROSEMARY J. REDFIELD

Program in Evolutionary Biology,
Canadian Institute for Advanced
Research and
Department of Biochemistry,
University of British Columbia,
Vancouver, British Columbia V6T 1Z3,
Canada

1. Kirkpatrick, M. & Ryan, M. J. *Nature* **350**, 33–38 (1991).
2. Goodgal, S. H. & Mitchell, M. A. *J. Bact.* **172**, 5924–5928 (1990).
3. Goodman, S. D. & Scocca, J. J. *Proc. natn. Acad. Sci. U.S.A.* **85**, 6982–6986 (1988).

SIR — Kirkpatrick and Ryan suggest¹ that there is growing support for the hypothesis that mating preferences evolve because of direct rather than indirect effects on female fitness. I believe that this conclusion is premature and stems, in part, from the way that tests have been attempted of one of the indirect effects, the Hamilton–Zuk² (or parasite) hypothesis.

Kirkpatrick and Ryan point out that many of the across-species comparative studies investigating the parasite hypothesis have found no relationship between the extremity of male display traits and the level of parasitism. They regard this as evidence that parasites are not important in the evolution of male displays. Indeed, they suggest that such comparative tests may be “the most fruitful way” to investigate the parasite hypothesis.

Previous authors^{3,4} have pointed out problems with various aspects of these tests, such as the difficulties associated with objectively measuring brightness and in obtaining holistic measures of the levels of parasitism. There seems, however, to be a more fundamental problem with the tests, stemming from the basic assumption that the parasite hypothesis predicts that the absolute level of parasitism in a species should be correlated with the extremity of the displays in that species. The crux of the parasite hypothesis is that the exaggerated male traits allow females to choose among males within their population and mate with a more resistant male. Thus, there is no reason to assume that there will be an across-species relationship between the level of parasitism and the degree of exaggeration of the trait.

Imagine, for example, a population of birds that has a very high level of parasitism but a very low variance in the level of parasitism or very low heritability in host resistance to the parasite. In this population, the parasite

hypothesis would predict the absence of exaggerated male display traits as females have very little to choose between males and, if heritability is low, nothing to gain by making a choice. Conversely, a population with low overall levels of parasitism but with high variance in parasite levels within the population and high heritability for resistance should have extreme display traits as females have much indirect fitness to gain by discriminating among males. So a more valuable across-species comparison would be to compare the extremity of the male trait with differences in the heritability and variance of resistance to parasites.

The relative importance of direct and indirect fitness effects in the evolution of exaggerated male traits can only be measured once these different factors have been more thoroughly examined in wild populations.

DAVID G. HASKELL

Section of Ecology and Systematics,
Division of Biological Sciences,
Cornell University,
Ithaca, New York 14853-2701, USA

1. Kirkpatrick, M. & Ryan, M. J. *Nature* **350**, 33–38 (1991).
2. Hamilton, W. D. & Zuk, M. *Science* **218**, 384–387 (1982).
3. Read, A. F. *Trends ecol. Evol.* **3**, 97–102 (1988).
4. Endler, J. A. & Lyles, A. M. *Trends ecol. Evol.* **4**, 246–248 (1989).

Simplified gene nomenclature

SIR — Vertebrate genes encoding several distinct voltage-gated potassium channels have recently been identified in many laboratories, but the names assigned to them

member of the *Shaker*-related subfamily 1 (1.1).

The tissue origin of these genes will not be included in the name as many of these genes can be expressed in several tissues. The species origin of a gene (such as Kv1.1) will not be included in the name, but could be described, for example, as rat Kv1.1 or

human Kv1.1. The list of genes will be updated periodically; genes published in the interim will have a ‘w’ prefix to signify ‘nomenclature being worked out’. Genes encoding channels whose voltage-dependence has not been experimentally confirmed can be indicated with an empty parenthesis, the ‘v’ to be added when functional data are available. As additional subfamilies are discovered, they will be numbered based on their date of discovery.

Other voltage-dependent potassium channel genes that are structurally distinct from the *Shaker/Shaw/Shab/Shal* proteins (for example the Isk channels) could be added to the current list of genes as a separate family (for example, Ks1.1 or Kvs1.1). The proposed nomenclature could be used for other types of K⁺ channels as well. Ligand-gated K⁺ channels could be similarly named, the particular ligand (Ca, ATP) being substituted for v.

A list of references and the affiliations of the signatories will be provided upon request from K.G.C..

K. G. CHANDY

Department of Physiology and
Biophysics,
University of California,
Irvine, California 92717, USA

Other signatories of this letter are: J. Douglas, G. A. Gutman, L. Jan, R. Joho, L. Kaczmarek, D. McKinnon, R. A. North, S. Numa, L. Philipson, A. B. Ribera, B. Rudy, L. Salkoff, R. Swanson, D. Steiner, M. Tanouye and B. L. Tempel. (Addresses available on request from K.G.C.).

A SIMPLIFIED NOMENCLATURE FOR A FAMILY OF VERTEBRATE VOLTAGE-DEPENDENT K⁺ CHANNEL GENES

Property	Xenopus	Mouse	Rat	Human
<i>Shaker</i> -related subfamily 1:				
Kv1.1	—	MBK1 MK1	RCK1 RBK1	HK1
Kv1.2	XSha2	MK2	RBK2 RCK5 NGK1	HK4
Kv1.3	—	MK3	RCK3 RGK5 KV3	HPCN3
Kv1.4	—	—	RCK4 RHK1	HK2 HPCN2 HK1
Kv1.5	—	—	KV1	HPCN1 HK2
Kv1.6	—	—	KV2 RCK2	HBK2
K(1).7	—	MK6 MK4	RK6	HaK6
<i>Shab</i> -related subfamily 2:				
Kv2.1	—	Mshab	DRK1	—
<i>Shaw</i> -related subfamily 3:				
Kv3.1	—	NGK2 Mshaw22	Kv4*	—
Kv3.2	—	Mshaw12	RKShIIIA Rshaw12	—
K(3).3	—	Kv3.3 Mshaw19	—	—
Kv3.4	—	Kv3.4	Raw3	—
<i>Shal</i> -related subfamily 4:				
Kv4.1	—	Mshal1	—	—
Kv4.2	—	—	RK5	—

*Kv4 is an alternatively spliced version of NGK2.

are frequently confusing. We propose a simplified nomenclature for one major family of these genes, based on sequence relatedness. This family comprises at least four subfamilies that represent vertebrate homologues of the four potassium channel genes in *Drosophila* (*Shaker*, *Shab*, *Shaw* and *Shal*). In the table a name is proposed for each gene. For example, the proposed name for MBK1/RCK1/RBK1 in our nomenclature is Kv1.1; that is, a K-channel gene (K) which is voltage dependent (v), and is the first identified

Conscious vs neural time

SIR — Conscious subjective experience and neural processes are phenomenologically independent; the first is accessible only to the subject individual, the latter to an external observer. The relationship between these phenomena can, however, be investigated. Our experimental analyses of the time-factors in this inter-relationship have provided evidence for at least two remarkable features; first, a substantial duration of continual cerebral neural activity, up to about 0.5 s, is required before a conscious experience appears (whether a sensation or intention to move)^{1,2}, and second, the subjective timing of the sensory experience appears

to have no significant delay. This timing appears to be subjectively referred back to an early signal arriving at the cortex within about 10–20 ms (ref. 3). Glynn argues⁴ that our proposal of subjective antedating is flawed, and that even if the proposal is sound, it does not violate any physical or physiological rules.

We had shown that the sensation of a skin stimulus pulse, applied well after the onset of a train of pulses to somatosensory cortex, was reported by subjects to appear before the sensation elicited by the cortical stimulus. It was empirically established that the cortical stimulus could not have elicited its sensation until at least 0.5 s after its onset; other evidence^{1,3} indicated that 'neuronal adequacy' for production of the skin-induced sensation also should have involved a similar delay.

The reported timing order for the skin versus cortically induced sensations therefore suggested a subjective antedating for the skin sensation but not for the cortically induced one; this proposal has been tested and confirmed³.

Glynn proposed⁴ an alternative explanation: to produce the conscious sensory experience, an extra latency of about 0.5 s would follow the adequate 0.5-s cortical stimulus but not the 0.5-s time for neuronal adequacy after the skin stimulus. But there is no experimental evidence that further cortical activity must follow the minimally adequate stimulus train applied to the cortex; indeed, electrophysiological recordings are negative¹. Consequently, adding an additional 0.5-s latency for the experience to appear only after the adequate cerebral stimulus is simply speculation. Glynn further proposes that no extra latency follows the empirically required minimum stimulus train applied to the medial lemniscus (ascending sensory pathway in midbrain). The striking absence of delay for subjective timing of lemniscal-induced sensation, in contrast to the cortically induced sensation³, is explained by our hypothesis without *ad hoc* assumptions of differences in latency.

The existence of a neural time-marker¹ does provide a neuronal event to act as a clock-time reference for subjective antedating, but it does not explain how that event is translated into an altered content of subjective timing. Everyone would agree that subjective experiences are intimately related to neuronal processes, but the relationships of subjective time (and other experiences) to neural activities are not definable, *a priori*, by physical observations or by the known physical laws; the relationships must be discovered^{1,2}. Glynn's views seem to rest on the unstated premise that a "neural-identity theory"⁵ explains the mind-brain relationship. But identity theory is a philosophical position, not a scientifically established one. In view of the special yet fundamentally important relation between the phenomenologies of conscious experience and the physical world, we need to keep an open mind about the possibilities for explaining it.

BENJAMIN LIBET

Department of Physiology,
University of California,
San Francisco,
California 94143-0444, USA

GLYNN REPLIES — By calling the period between the completion of a cortical stimulus and the evoked sensation an "extra latency" — as if there had already been a latent period — and the period between the completion of a skin stimulus and the evoked sensation a "time for neuronal adequacy", Libet makes my suggestion that the two latent periods are similar⁴ look like a suggestion that they are very different. His letter also misrepresents my suggestion about the latent period that follows stimulation of the medial lemniscus.

Optimal circular packing

SIR — In his News and Views article¹, Ian Stewart recounts the amusing story of how the clathrin cages of coated vesicles beat the mathematicians at solving the problem of finding optimal coverings of spheres by circles. I would like to draw attention to another class of structures, the inorganic clathrate hydrates, that may have the same mathematical ability.

The structures of the clathrate hydrates² are based on a space-filling packing of polyhedra with pentagonal and hexagonal faces. An example is provided by the structure of the hydrate of chloroform (sometimes called the type II clathrate hydrate structure), which has been known for some time³. As well as the pentagonal dodecahedra which solve the 12-circle problem, the structure contains the hexakaidecahedra with 12 pentagonal and 4 hexagonal faces (*b* in Stewart's article), that have now been shown⁴ to correspond to the best-known solution to the problem of the optimal sphere covering by 16 circles.

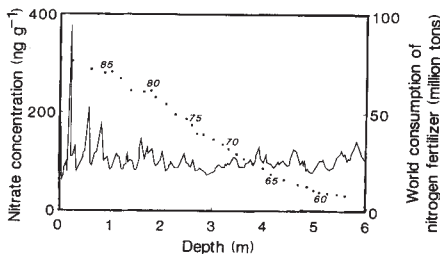
The dual of the hexakaidecahedron, where now the vertices represent the 16 circle centres, has an even longer history⁵ in crystal chemistry, where it is called the Friauf polyhedron. It may be derived by capping the hexagonal faces of a truncated tetrahedron and is a conspicuous feature of many alloy structures². However, in this context it probably arises as an attempt to solve the complementary circle packing (rather than covering) problem.

MICHAEL O'KEEFFE

Department of Chemistry,
Arizona State University,
Tempe, Arizona 85287,
USA

Fertilizer and snowfall nitrate

SIR — Turner compares¹ global consumption of nitrogen fertilizer from 1950 to 1987 with measurements of the nitrate concentration in snowfall at the South Pole (station T) taken from Mayewski *et al.*². Unfortunately, the data from Mayewski's original paper (Fig. 3 of ref. 2) have been misunderstood. The data reproduced by Turner as nitrate in snowfall



Nitrate concentration (solid line) in 6 metres of snowfall at South Pole station T (ref. 1) compared to world consumption of nitrogen fertilizer (dotted line) (from ref 2).

are in fact the October total ozone concentration from Halley station, Antarctica. A diagram showing the correct comparison is produced above. Turner states that the apparent rise in nitrate is a linear trend with a 60 per cent increase between 1975 and 1985, and compares this to a 62 per cent increase in global consumption of nitrogen fertilizer over the same period. The snowfall-nitrate data, however, appear essentially flat, apart from unusually large peaks in the 1987 and 1988 snowfall. It seems unlikely on this evidence that a correlation exists between world consumption of nitrogen fertilizer and nitrate deposition at the South Pole.

R. MULVANEY

British Antarctic Survey,
Natural Environment Research Council,
Madingley Road,
Cambridge CB3 0ET, UK

1. Stewart, I. *Nature* **351**, 103 (1991).

2. Wells, A. F. *Structural Inorganic Chemistry* 5th edn (Clarendon, Oxford, 1984).

3. Clausen, W. F. *J. chem. Phys.* **19**, 259–260 (1951).

4. Tarnai, T. *J. molec. Biol.* **103**, 485–488 (1991).

5. Friauf, J. B. *J. Am. chem. Soc.* **49**, 3107–3114 (1927).

1. Turner, R. E. *Nature* **349**, 469–470 (1991).

2. Mayewski, P. A. & Legrand, M. R. *Nature* **346**, 258–260 (1990).

The basic difficulty, however, is straightforward. As Churchland⁴ first pointed out, Libet's argument for "an automatic subjective referral of the conscious experience backwards in time"³ depends crucially on his assumption that the sensation following a train of (individually subthreshold) shocks to the cortex starts immediately, or almost immediately, after the shock that makes the train long enough to be effective; in other words, that there is only a very short latent period. The arguments that Libet used⁵ in reply to Churchland's criticisms, arguments which attempted to show that the latent period must be extremely short, are themselves unsatisfactory⁴. Because we do not know the duration of the latent period, any assumption about its length must be arbitrary, though not all assumptions are equally plausible. Libet is correct in pointing out that if the latent period is indeed very short, his results appear to lead to a surprising and exciting, though not impossible, conclusion. But until the assumption is supported by evidence, scepticism is more advisable — if less enjoyable — than excitement.

IAN M. GLYNN

Physiological Laboratory,
University of Cambridge,
Downing Street,
Cambridge CB2 3EG, UK

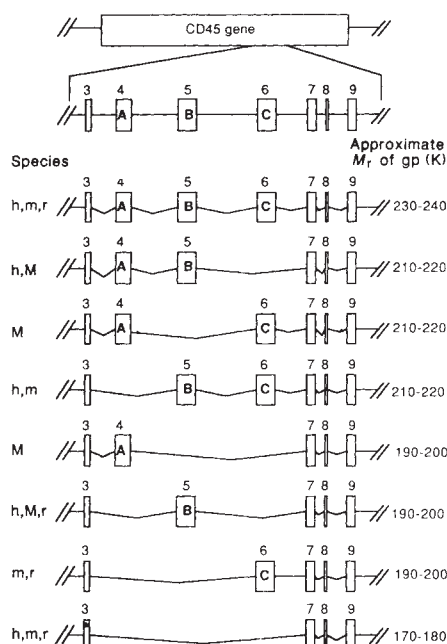
1. Libet, B. in *Models of Brain Function* (ed. Cotterill, R. M. J.) 35–49 (Cambridge Univ. Press, 1989).
2. Libet, B. *Behav. Brain Sci.* **8**, 529–566 (1985).
3. Libet, B., Wright, E. W. Jr, Feinstein, B. & Pearl, D. K. *Brain* **102**, 193–224 (1979).
4. Glynn, I. *Nature* **348**, 477–479 (1990).
5. Feigl, H. & Pepper, S. C. in *Dimensions of Mind* (ed. Hook, S.) 24–56 (New York Univ. Press, 1960).
6. Churchland, P. S. *Philosoph. Sci.* **48**, 165–181 (1981).
7. Libet, B. *Philosoph. Sci.* **48**, 182–197 (1981).

Confusion over CD45 isoform

SIR — Bell and Sparshott report¹ the apparent interconversion of rat T-lymphocyte populations expressing distinct isoforms of CD45, the leukocyte common antigen. Their extrapolations of the data presented are disturbing, however.

CD45 isoforms are generated from a single gene by the alternative use of three variable exons termed A, B and C (corresponding to exons 4 to 6 of the CD45 gene, see figure)^{2–5}, and isoform-specific CD45 antibodies are classified according to the exon on which their reactivity depends. For example, CD45RA reagents bind epitopes encoded by exon A on several high molecular mass CD45 isoforms, and CD45R0 indicates specificity for the smallest isoform, which lacks sequences encoded by exons A, B and C (ref. 6).

In humans, CD45RA and CD45R0 define largely reciprocal T-cell populations⁷. When activated *in vitro*, CD45RA T cells acquire CD45R0 (ref. 7). Frequencies of T-cell memory responses are high among CD45R0, but low among CD45RA populations, suggesting that the stable acquisition of



Schematic representation of the CD45 gene (based on refs 2–5, 12, 14). A, B and C, variable exons (4 to 6) as assigned by Streuli *et al.*⁵; h, m and r, human, mouse and rat cloned cDNAs, respectively; M, murine isoforms demonstrated by polymerase chain reaction¹¹; numbers on the right, relative molecular mass (M_r) for translated and glycosylated products¹⁴.

CD45R0 determinants may accompany human T-cell priming *in vivo* as well as *in vitro*^{8–10}. Bell and Sparshott propose a revision of this hypothesis¹. In their study, rat T cells are enriched or depleted for CD45RB expression with the CD45RB reagent Ox22 (ref. 11), injected into nude (T-cell deficient) recipients, and allowed to expand for several weeks or months. CD45RB-enriched, as well as CD45RB-depleted, populations give rise to progeny of mixed CD45RB positive/negative phenotype. This finding, the authors conclude, is incompatible with the stable acquisition of CD45R0 during human T-cell memory formation⁷. The apparent conflict hinges on the assumption that CD45RB negative (Ox22[−]) T cells necessarily express CD45R0 (ref. 1). A review of the possible CD45 isoforms lacking exon B-encoded sequences (see figure) shows that this equation is not justified: three variable exons allow eight permutations, at least six of which have been isolated as complementary cDNA clones^{2–5}. Confirmation of all eight isoforms comes from studies using the polymerase chain reaction with exon-specific primers¹². At least one cloned rat CD45 cDNA lacks exon B but contains exon C (ref. 3), and immunoprecipitation analysis indeed indicates that not all high molecular mass rat CD45 species are recognized by Ox22 (ref. 11). CD45RB-positive as well as CD45RB-negative rat T-cell populations are therefore heterogeneous in terms of their CD45 isoform expression, and conclusions regarding the gain or loss of CD45R0 cannot be drawn

from the experiments reported by Bell and Sparshott¹.

Extrapolation of the data is also tenuous because, in humans, unlike rats, CD45RB is expressed by the great majority of T lymphocytes¹³ and (in contrast to CD45RA and CD45R0) has not been established as a marker for T-cell heterogeneity.

The generation of antibodies specific for rat CD45RA and CD45R0 might clarify the situation. Alternatively, methods such as the polymerase chain reaction could be employed to characterize CD45 isoform expression by rat T-cell populations. But more general obstacles complicate the analysis of progeny derived from large, heterogeneous precursor populations. To rule out the possibility that contaminating cells are responsible for their results, Bell and Sparshott inject mixtures of allotype-marked T cells, enriched or depleted for CD45RB expression. They find CD45RB-positive and negative progeny of both allotypes, and conclude that all injected cells expand equally well. The data would also (and perhaps more in line with other haematopoietic reconstitution models) be explained by the presence of specific precursors giving rise to heterogeneous progeny of both allotypes. Donor-derived cells undergo extensive initial proliferation, but T-cell numbers reach relatively stable levels some time (as judged by the relative expansion reported after 3 weeks and 7.5 months). The chimaeras are therefore able to regulate their peripheral T-cell pool, and homeostatic mechanisms could also explain stable ratios between CD45RB-positive/negative T lymphocytes.

If these problems can be resolved, the approach taken by Bell and Sparshott may bring a constructive revision of current concepts regarding CD45 isoform expression during T-cell ontogeny. As it stands, such conclusions seem premature.

MATTHIAS MERKENSCHLAGER

Department of Immunology,
Institut de Chimie Biologique,
11 rue Humann,
67085 Strasbourg Cedex, France

1. Bell, E. B. & Sparshott, S. M. *Nature* **348**, 163–166 (1990).
2. Ralph S. J., Thomas M. L., Morton C. C. & Trowbridge I. S. *EMBO J.* **6**, 1251–1257 (1987).
3. Barclay, A. N., Jackson, D. I., Willis, A. C. & Williams, A. F. *EMBO J.* **6**, 1259–1264 (1987).
4. Saga Y., Tung J.-S., Shen F.-W. & Boyse, E. A. *Proc. natn. Acad. Sci. U.S.A.* **83**, 6940–6944 (1986).
5. Streuli, M., Hall, L. R., Saga, Y., Schlossman, S. F. & Saito, H. *J. exp. Med.* **166**, 1548–1566 (1987).
6. Knapp, W. *et al.* (eds) *Leukocyte Typing IV. White Cell Differentiation Antigens* (Oxford University Press, 1990).
7. Terry, L., Pickford, A. & Beverley, P. C. L. in *Leukocyte Typing III. White Cell Differentiation Antigens* (eds McMichael, A. J. *et al.*) 225–227 (Oxford University Press, 1987).
8. Beverley, P. C. L. *Immun. Lett.* **14**, 263–267 (1987).
9. Merkschlager, M., Terry, L., Edwards, R. & Beverley, P. C. L. *Eur. J. Immun.* **18**, 1653–1658 (1988).
10. Merkschlager, M. & Beverley, P. C. L. *Int. Immun.* **1**, 450–459 (1989).
11. Spickett, G. P., Brandon, M. R., Mason, D. W., Williams, A. F. & Woollett, G. R. *J. exp. Med.* **158**, 795–810 (1983).
12. Chang, H.-L., Zaroukian, M. H. & Esselman, W. J. *J. Immun.* **143**, 315–321 (1989).
13. Pulido, R., Cebrian, M., Acevedo, A., Landazuri, A., and Sanchez-Madrid, F. *J. Immun.* **140**, 3815–3827 (1988).
14. Thomas, M. L. *Rev. Immun.* **7**, 339–369 (1989).

Hook, line and sinker

John Shepherd

Crisis in the World's Fisheries: People, Problems & Policies. By James R. McGoodwin. Stanford University Press: 1990. Pp.235. \$35.

THERE are problems affecting fisheries all around the world, and they have been with us for a very long time. A diligent historian could find examples among the earliest written records: certainly the Romans were troubled about the morality of catching dolphins. In *The Provident Sea* (Cambridge University Press, 1988), David Cushing quotes the Roman scholar Oppian, writing on this subject in the second century AD, and gives many other interesting historical references (and some delightful illustrations). The problems are often due to sudden fluctuations of stock size, or conflicts of interest between different groups of fishermen, often triggered by the introduction of a new and more efficient form of fishing gear — there were special restrictions placed on beam trawling in England as early as 1376. Concern over the effects of pollution and the depletion of stocks by over-fishing has intensified with industrialization, but has been with us since the middle of the nineteenth century. The International Council for the Exploration of the Sea (ICES) was specifically set up because of such concern, and will be 90 years old next year.

I therefore turned to James McGoodwin's book with some trepidation. What new problem was now upon us, to make life even more difficult for those who seek to manage fisheries? I suppose I should be thankful that I failed to find anything novel: the crisis seems to be more in the eye of the publisher's marketing department than anywhere else. The book is in fact a fairly wide-ranging review of the social aspects of small-scale (artisanal) fisheries; the author is a professor of anthropology at the University of Colorado. The coverage is exclusively marine, even though Colorado is a long way from the sea. The acknowledgements suggest that the book is the fruit of time spent at the Woods Hole Oceanographic Institution in New England. The scope is, however, global, with many examples from the Americas, Asia and the Pacific. Given the emphasis on small-scale fisheries, the focus on developing countries is understandable, even though such fisheries are common in developed countries too.

The account of the history of the development of the science of fisheries management is brief but reasonably accurate. Even though the focus is on small-scale fisheries, it is surprising to find almost no discussion of the international dimension of the problems associated with fisheries. There is no mention of how difficult even supposedly



Harvest from the deep — unloading tuna in the Philippines.

advanced nations find it to resolve such problems, as illustrated by the dispute between the United States and Canada over Georges Bank, or the continuing travails of the European Common Fisheries Policy. The author commends management systems that give property rights to the participants, but makes no mention of current attempts to do just that, for example, in Iceland, Canada and New Zealand. There is, in fact, no mention of individual transferable quotas, or individual vessel effort limits as currently used in the Netherlands. Is experience

gained in large-scale fisheries really of so little relevance?

McGoodwin persistently suggests that the professionals involved in fisheries management ignore and even belittle the social aspects of the problem. He also attacks the classical bioeconomic analysis of fisheries, which asserts that in an unregulated fishery, fishing effort tends to expand to a point where the average participant can only just earn enough to cover his short-run operating costs. In economic jargon, the potential 'rent' from the fishery is dissipated. I can find no basis for his claim that this principle is fatally flawed, except that the argument is usually made on the assumption of common property rights, whereas it should more correctly rely on the existence of open access.

Impact Photos

The criticism directed at the biologists and economists who are deemed to have led us all astray is followed by a chapter entitled 'A New Era in Fisheries', which contains a section with the even more portentous subtitle 'The Dark Before the Dawn'. This actually refers to the general extension of fishery limits to 200 miles during the 1960s and 1970s, but culminates in the suggestion that, until now, fisheries management has ignored the social aspects of fishing, and the mechanisms used by indigenous peoples to regulate fishing, and that if we would only take note of these, all would be well.

This, I fear, is naive. And the examples given in the chapters that follow all amount to either limiting access, or limiting efficiency, in one way or another. Does McGoodwin really believe that the professionals he criticizes are not aware that these may be useful tools, and that they have not attempted to apply them many times already? Perhaps he should have spent more time talking to the staff of the National Marine Fisheries Service down the road in Woods Hole. They could have brought him up to date on the problems now faced by the New England Fisheries Management Council,

a body in which the fishermen's representatives have a dominant role, precisely as he suggests for the future. The council has not been a success, for the simple reason that, like athletic training, fisheries management only does any good if it hurts.

A more balanced discussion of the possible limits to cooperative management would have been welcome. Is there any validity in the idea that cooperation only works properly if all the participants know one another and, preferably, are related to one another? Or any truth in the notion that

cooperation is liable to break down when threatened by outsiders, especially if they speak another language? McGoodwin himself says "Local fishers will indeed stop fishing with restraint when newcomers invade their property." If this is so, how do we construct cooperative management systems that will stand the test of time?

It is regrettable that McGoodwin makes such a point of condemning the attempts made so far by fisheries' managers to find solutions to the perennial and intractable problem of the rational and equitable exploitation of fisheries. It leaves one ill-disposed towards the remainder of his thesis, especially as this is presented with a gloss of New Age optimism, which implies that here at last is the missing ingredient we have all been seeking. In fact, what he proposes is that attempts to regulate fisheries should take more account of the social problems of those involved (mainly the fishermen), should try to work through cooperation and self-regulation, and should involve the fishermen

more intimately in the setting of regulations and the formulation of policies. This is not new, indeed ICES has for many years been organizing so-called 'dialogue meetings' of fishermen's representatives, biologists, economists and administrators, precisely to try to make progress along these lines, with only modest success. Nevertheless, most people involved agree that we need to move in this direction, and another reminder would do no harm. Unfortunately, this reminder will annoy the converted, and is unlikely to persuade the rest.

McGoodwin bemoans the fact that fisheries management agencies employ so few sociologists and anthropologists. If *Crisis in the World's Fisheries* is a fair example of what they can do for us, I doubt that that will change. □

John Shepherd is at the Ministry of Agriculture, Fisheries and Food, Fisheries Laboratory, Pakefield Road, Lowestoft, Suffolk NR33 0HT, UK.

Language and lexicography

Paul Fletcher

The Science of Words. By George Miller. W. H. Freeman & Co.: 1991. Pp.276. £10.95, \$12.95.

In the popular view, the science of words is enshrined in lexicography. And the rich tradition of dictionary-making in English, from before Dr Johnson to Dr Burghfield, is acknowledged not only at various places in George Miller's text, but in his current project, WordNet. This involves the development of a dictionary of English, accessible on-line, whose lexical database is organized on the basis of current psycholinguistic theories.

Conventional dictionaries provide information primarily about the meaning of a word, along with its pronunciation and spelling, and the syntactic category to which it belongs. This, as Miller points out, is a good start. Any theory of the mental lexicon, of how the individual organizes and accesses his store of words in memory, will begin from this form-meaning association. But it must also take account of the networks of relationships, of both form and meaning, that relate individual words. Speakers know, or learn, that a pit-bull terrier is a kind of dog, that leaping is a particular way of jumping, that tall people are at the upper end of a scale of height and short people at the lower end, that there is a similarity in meaning between 'book' and 'monograph'. In addition to these meaning relationships (respectively hyponymy, troponymy, antonymy and synonymy), speakers also know about relationships of form. Writers write, and researchers research, but porters do not port.

Parliament recesses, but a recession is something different, even though there does seem to be a meaning-preserving derivational relationship between the verb 'process' and the noun 'procession'.

It is Miller's career-long involvement with the representations of words and their form and meaning relationships in the minds of speakers that informs and shapes this book. He systematically examines the three-sided character of words identified by the lexicographer — spoken/written form, syntactic role and concept — from the perspective of the mental lexicon. As is appropriate for a general, introductory text, this involves some exposition of the development and current status of writing systems, of the physics of speech and the structure of phonologies in languages, and of how the syntax of sentences is organized. These accounts of how words are pronounced or written, and how they are organized into permissible linear sequences are never less than lucid, elegantly written and, most remarkable to someone familiar with the standard, visually drab linguistics text, lavishly and imaginatively illustrated. But these sections serve as background for Miller's main interest, the words themselves. How words evolve, how many we know, how they are learned, the nature of their storage and retrieval, lexical loss through brain damage, the boundaries between linguistic and encyclopaedic knowledge are some of the topics that Miller deals with. In a synthesis of his own work (for he is a contributor of the first rank to this field) and other research, he writes in a style which is clear, while eschewing superficiality.

Some of the most interesting current work on the mental lexicon concerns its acquisition by children. Miller enjoins us to be sceptical about estimates of vocabulary size but, that done, suggests a possible vocabulary size for the average US high-school graduate of 60,000 root words. Assuming these are

acquired over 16 years, this suggests (if we allow another assumption, of a constant rate of acquisition) something over ten new words a day. Because, at least in the early stages of language development, the only information available to the pre-literate child about the form of a word is through the auditory channel, we need to consider an acquisition device that can assemble the sound-meaning relationships embodied in words with some facility.

Recent work, reviewed by Miller, shows that there is, for the learning of a word, an initial 'fast mapping' phase. In one experiment, three-year-old children were presented just once with the presumably unfamiliar word 'chromium' to refer to an object that was olive-green. A week later, when tested, half of them could recognize the word (that is, had remembered its phonetic shape) and knew it was a colour word. They could assign the unfamiliar term to the semantic field of colour names, even though they were unable to identify the colour precisely. Following the initial acquisition of phonetic form and some meaning information, there is a slower second phase, which may require many more instances of the particular word for its syntactic role and full meaning to become clear for the child. Different syntactic categories have been investigated, and it appears that for fast mapping, verbs cause children more difficulties than nouns. Although Miller does not deal directly with this issue, reasons for the discrepancy between nouns and verbs in acquisition can be derived from a discourse on semantic relations between verbs, which is unusual for a general text in being quite original.

For a general text on the science of words and their users, which can be said to be introductory but which is also lucid, interesting and original, look no further. It deserves to be popular. □

Paul Fletcher is in the Department of Linguistic Science, University of Reading, PO Box 218, Reading RG6 2AA, UK.

New in paperback

■ *Cosmos: The Growth of Order in the Universe* by David Layzer provides a highly individualistic analysis of cosmology, starting with subatomic particles and ending up with the evolution of the human mind. Published by Oxford University Press, price is £8.95. □

■ Many scientists have difficulties deciding how best to present their findings for the purposes of publication; *Writing Successfully in Science* by Maeve O'Connor should reduce the pain. Published by Harper Collins at £8.95 (also available in hardback, price £25). □

■ All those with an interest in ornithology or the evolution of animal behaviour should find something to their liking in *Behavior and Evolution of Birds*, edited by Douglas Mock. The book is a collection of articles from *Scientific American* and is published by Freeman at \$12.95. □

Stormwatch

Philippa Lloyd

Watching the World's Weather. By W. J. Burroughs. Cambridge University Press: 1991. Pp.196. £17.50, \$24.95.

SATELLITE images of the weather can be seen every day on our television screens, but what can be learned from them? On a superficial level, we are able to watch storms develop, track hurricanes (clearly identifiable by their catherine-wheel-like structure), and even perceive the plumes of smoke from burning oil fields. But how do forecasters use satellites to predict the weather, and has their arrival significantly improved our weather forecasts? Burroughs sets out to address these questions and others by providing an interesting, but rather pedagogical, account of the development of the field of satellite meteorology, together with an appraisal of the role of satellites in studies of climate change, and a considered look at the prospects for the future.

For those not familiar with the physical processes that are responsible for generating the weather patterns, Burroughs provides a succinct introduction. The underlying principle is that the Earth must be in radiative balance with the Sun — simply put, this means that the amount of solar radiation absorbed by the atmosphere and surface of the Earth must equal the amount of terrestrial radiation emitted — and for this balance to be achieved, large amounts of energy have to be transported through the atmosphere. So, armed with a basic understanding of the “global weather machine”, both the science and technology aspects of satellite meteorology are then described in detail and can be easily understood.

The first photographs of weather systems were obtained in 1947 from a camera strapped to a V2 rocket, and by 1960 the first operational satellite was in orbit. The images transmitted to Earth were of immediate use in providing images of the developing weather, but to improve forecasts significantly, satellite instruments had to be able to provide accurate and reliable measurements of temperature, humidity, pressure, rainfall and wind speed.

As Burroughs explains, and indeed as one might imagine, determining the temperature of the atmosphere is actually quite a complicated business, as temperature does not vary with height in a simple way: in the lower atmosphere (the troposphere), temperature decreases approximately linearly with height, because energy is transferred mainly by convective processes, whereas immediately above the troposphere, radiative processes dominate and the temperature remains more or less constant with height. Above about 100 km, temperature starts to increase with height. This is a very simplistic description of the temperature profile of the atmosphere, because clouds, weather systems, extremely cold regions on the Earth's surface (the

valuable information about a certain section or slice of the atmosphere. To retrieve a complete temperature profile, a number of different wavelengths must be used, preferably in both the infrared and microwave regions of the spectrum. A similar principle is adopted to retrieve the humidity (water vapour) profile of the atmosphere.

In contrast, to deduce wind speed, wave height and the strength of ocean currents, for instance, radar systems are used, which send out pulsed beams of microwave radiation. By comparing the characteristics of the reflected pulse with those of the incident beam, estimates of these properties can be made.

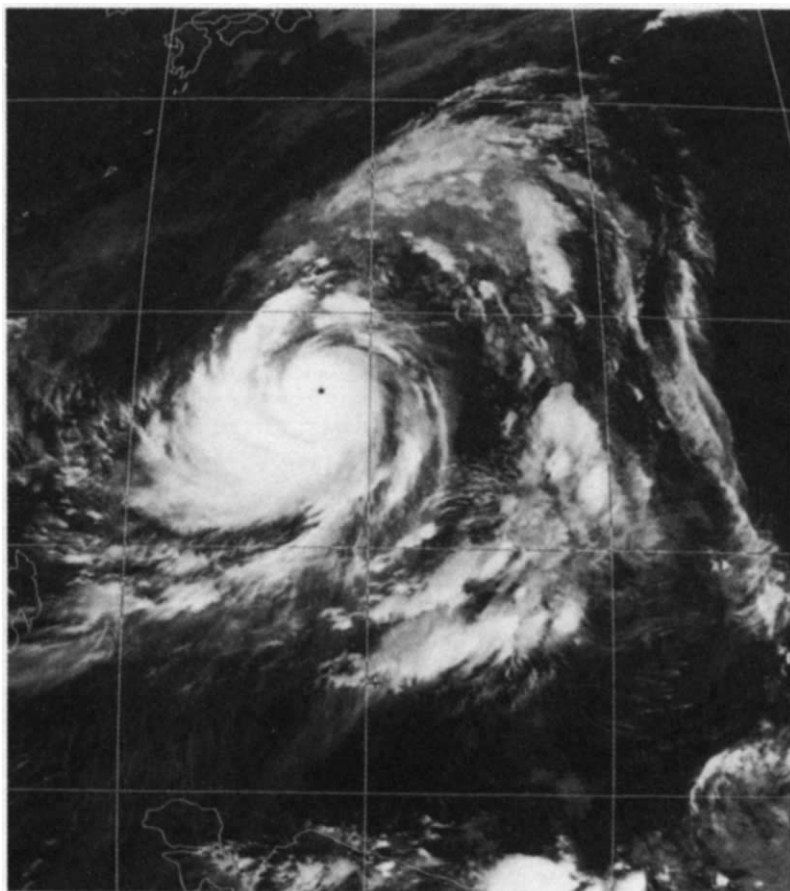
As well as measuring the upwelling radiation, and hence determining the temperature and humidity profiles, some satellite instruments have been designed to examine the properties of the sunlight reflected from the surface of the Earth, allowing estimates to be made of the biological productivity of the oceans, the extent of different types of vegetation, and the areal coverage of sea ice and snow. Satellites can therefore be used to monitor both local and global changes in the environment over a period of years — an invaluable capability for those interested in examining the climate response to increasing concentrations of gases such as carbon dioxide and methane.

The potential value of satellite observations is clear, but there have been considerable problems in exploiting the data to the full in weather forecasting, as Burroughs makes clear. To analyse the data and combine the satellite measurements with ground-based observations has put computer science to the

test. Indeed, only relatively recently, with the advent of more powerful computers and more sophisticated analytical techniques, have satellite data started to make their presence felt.

Watching the World's Weather is certainly informative, and honest about the limitations of satellite data as well as the undoubted benefits. The occasionally rather dry prose is leavened by the inclusion of some spectacular satellite pictures of storms and hurricanes. The satellite image on the television screen will never seem quite the same again. □

Philippa Lloyd is an Assistant Editor of *Nature*.



Typhoon Tip — the most intense typhoon ever measured in the Pacific. This infrared image shows the clearly defined 'eye' of the storm.

Arctic and Antarctic, for example) all serve to perturb the profile; but the complexity of measuring the temperature should at least be clear. By making clever use of the fact that the amount of energy radiated by gases in the atmosphere is a function of their temperature, instruments have been designed to measure the upwelling radiation received at the satellite over a range of wavelengths. If the atmosphere absorbs strongly at a particular wavelength, then most of the radiation measured will come from the higher levels of the atmosphere. Correspondingly, if the atmosphere absorbs weakly, then the radiation received comes from near the Earth's surface. Hence each wavelength used yields

Preserving species

Mark Williamson

Nature Reserves: Island Theory and Conservation Practice. By Craig L. Shafer. *Smithsonian Institution Press*: 1991. Pp.183. Hbk \$39.95, £31.25; pbk \$15.95, £12.50.

CONSERVATION can be a depressing subject. The steady increase both of the human population and of economic activity leads to the loss of habitats and so to the loss of species. The most obvious defence is to create reserves. Many questions arise: how many; where; what size and shape; should there be buffer zones; can there be corridors for wildlife connecting reserves?

These questions are addressed in this book, as is the larger question: does ecological theory help to find better answers? All reserves are, to some extent, comparable to islands, so it is not surprising that 'island theory', more accurately the MacArthur-Wilson theory, has been much discussed, and here quite critically.

Pictures and diagrams are included that bring home the immensity of the problem. Along with familiar maps of the reduction in the past few centuries of forest in Wisconsin, United States, and São Paulo, Brazil, and of heathland in Dorset, England, there are satellite photographs of parts of the contiguous United States showing how small many reserves are as a proportion both of the landscape and of the original ecosystem they represent. These photographs also show how these reserves are bound in by urban, agricultural and forest-extraction developments. The fragmentation, isolation and small size of the residua, the reserves, are all too evident.

The intention of this nicely produced book is given by some quotations: "a discussion of the theoretical aspects of nature reserve size, isolation, and design, as well as planning and management implications", "It is no longer wise for planners and managers simply to rely on any one scientist's translation of this literature into conservation guidance", "I have tried to include conflicting interpretations on as many points as possible" and "brought in aspects of the real world of nature reserve management, mostly ignored up to now in such papers".

This attempt to see all sides of the question in less than 200 pages is remarkably successful. There is a lot of 'A says this but B says that'. There are irritations like a 1966 date for a dataset that is only statistically reliable from 1969; the claim that H. C. Watson's important 1835 book (*Geographical Distribution of British Plants*, Longman) was not published, and an index that, for authors, refers only to the frequently quoted ones. More seriously, although there is much on extinction there is very little on invasion or

immigration. The fashionable, if largely misleading, topic of 'metapopulation' is not mentioned, though it is based on a concept of an archipelago of populations. Along the way, there is much useful discussion of species-area relationships, minimum population size, the rate of species loss, as well as of the design of reserves, all of which should be known to managers and administrators. The difficulty that such people will have is that, although the broad concepts and guidelines are clear, rather few detailed decisions can follow directly from them.

The final conclusions are somewhat bland, but this is a proper reflection of the many and deep disagreements about the validity and bearing of much of the theory. Here are six from a list of 26: "The more land you set aside, the more species you will preserve"; "Habitat fragmentation and nature reserve insularization should be discouraged"; "Small populations should be avoided"; "To maintain biological diversity, laissez-faire management of nature reserves should be the exception, not the rule"; and, "Regional planning of nature reserves must take human population growth into account and consider the social and economic conditions on adjacent lands."

In that last conclusion are the major difficulties. How do you raise the money to set aside land for reserves and manage them properly? How do you get planners and

executives to take decisions that assume the reserves preserve what they are meant to? In Britain, the questions have not usually been on the design of reserves; they have been on acquiring land, or making agreements on land, or on preventing development. The political, economic and social aspects of these difficulties are not dealt with. They vary much from country to country, but in the end they will determine which species survive.

At the moment, we can see no prospect of the end of population growth and changes in land use. Shafer boldly treats, from a palaeobiological perspective, the problem of evolutionary arrest. If planning and acquisition have a timescale of ten years, and management of 100, evolution mostly comes in at a scale of 10,000 years upwards. In the scales in between come the problems caused by climatic and other environmental change, and they are not mentioned at all.

On the whole, this is a useful work of reference. Those who most appreciate the uncertainties of science will be best placed to argue convincingly in the public arena. □

Mark Williamson is in the Department of Biology, University of York, York YO1 5DD, UK.

■ Of related interest is *The Scientific Management of Temperate Communities for Conservation*, edited by I. F. Spellerberg, F. B. Goldsmith and M. G. Morris. Published by Blackwell at £45 (Hbk); £26.50 (Pbk). □

Inside story

John Gittus

The Truth About Chernobyl. By Grigori Medvedev. *Tauris/Basic Books*: 1991. Pp.252. £14.95, \$22.95.

THIS year marks the fifth anniversary of the nuclear accident at Chernobyl and there has been a great deal of publicity for the plight of those still living on land contaminated by radioactive material that escaped from the reactor. Medvedev was chief engineer when the Chernobyl reactor was under construction, and was critical of the design at that time. Indeed, a number of serious criticisms had been published in Russian journals before the accident, showing the depth of concern felt by engineers living in a society that in those days was anything but open (*Glasnost* stemmed from the Chernobyl epoch).

Medvedev's technical account of the things that happened during the accident accords generally with our understanding, but his book certainly grips the reader's attention, bringing to life the people involved in a way that no dry, technical account could ever do. Most readers will find it difficult to follow the details of how the accident occurred from the account given in the book, but this does not detract from the tension that the author generates as he skillfully pieces together the sorry catalogue of events that led to prompt criticality and meltdown.

Our understanding had been that the contributing faults in the design of the reactor were its positive reactivity coefficients, which meant that it could not safely be left to 'simmer' without running the risk of boiling dry — exactly what happened in the accident — and 'positive scram'. Medvedev confirms this, pointing out that positive scram, the rise in reactivity that occurred when the operators tried to reduce reactivity by scrambling (or switching off) the reactor, was not implicated by the Soviets until some time after the accident. A cover-up or just ignorance? Of particular interest to me is the graphic account given of the brainstorming session that led to the decision to use helicopters to drop thousands of tons of sand on the open, burning reactor, because this was something that I suggested in a number of radio and television interviews just after news of the accident reached us in Britain.

Chernobyl was just about the worst nuclear-power accident imaginable, and this book is a telling account of how it occurred and of its immediate consequences. The account of exactly how the accident was triggered, when deciphered, does not add to what we know already, but a dramatic human dimension is given by the verbatim testimony of operators, emergency workers, inhabitants of the neighbouring town, party officials and senior people who descended from Moscow. □

John Gittus is at the British Nuclear Forum, 22 Buckingham Gate, London SE1E 6LB, UK.

Recent variations in Arctic and Antarctic sea-ice covers

Per Gloersen & William J. Campbell*

Laboratory for Hydrospheric Processes, NASA Goddard Space Flight Center, Code 971, Greenbelt, Maryland 20771, USA

* Ice and Climate Project, US Geological Survey, University of Puget Sound, Tacoma, Washington 98416, USA

Variations in the extents of sea-ice cover at the poles and the areas of open water enclosed within them were observed every other day during the interval 1978–1987 by a satellite-borne scanning multispectral microwave radiometer. A band-limited regression technique shows that the trends in coverage of the Arctic and Antarctic sea-ice packs are not the same. During these nine years, there are significant decreases in ice extent and open-water areas within the ice cover in the Arctic, whereas in the Antarctic, there are no significant trends.

THE scanning multichannel microwave radiometer (SMMR) which operated on board the Nimbus-7 satellite from October 1978 to August 1987, the longest interval of observation yet accomplished by a satellite-borne passive microwave instrument, allowed observations to be made of the entire Arctic and Antarctic sea-ice covers every two days through the clouds during night and day. It has thus provided a unique almost decade-long record of the large-scale behaviour of sea ice on Earth.

Sea ice is an important part of the global climate system. The two key aspects of the complex and dynamic sea-ice covers are ice extents (the total area enclosed by the ice–ocean edges) and the area of open water within them, consisting of cracks called leads and large amorphous openings called polynyas, from the Russian word for lake. The sea-ice extent affects the amount of solar radiation absorbed in its hemisphere and alters the atmosphere–ocean exchanges of heat, moisture and momentum. The area of open water within the ice pack, even when small compared with the ice extent, is critically important because the fluxes of heat and moisture through a given area of open water are 2–3 orders of magnitude greater than through the same area of sea ice¹. The multispectral and dual-polarized observations with the SMMR allow determinations of ice extent and open water with greater accuracies than those obtainable with earlier satellite sensors. The accuracy of the ice edge positions determined by SMMR is about half the diameter of the SMMR field of view, or about 25 km (ref. 2). The precision of the open-water estimates is estimated to be within 5% in the Arctic Ocean and 9% for other seas where large amounts of thin ice are present³.

The SMMR observations made over nearly nine years, uncorrected for instrument drift, were used in a recent paper⁴ to derive variations of the extents of Arctic and Antarctic sea ice and the open-water areas within the ice–ocean edge. The ordinary linear regression analysis of these uncorrected data indicated that trends of –4% and –5% in the maxima and minima, respectively, of the annual ice extents occurred in the Arctic over the 8.8-year span. The goodness-of-fit parameters (R^2) were 0.4 and 0.1, respectively. In the Antarctic, the trend in the ice-extent maxima was –3% ($R^2 = 0.1$). But a statistically more significant negative trend of 5% ($R^2 = 0.8$) was found for the peaks in the sum of the two polar ice packs, which is termed the annual global ice extent. These trends were obtained from a very small sample of the total SMMR data set: that is, only the data at the time of

the peaks of the winter extents and the troughs of the summer extents.

The key question posed by this earlier paper, and left unanswered, was the following: if the maxima of the global ice extents show a significant negative 8.8-year trend, what occurred within the separate Arctic and Antarctic sea-ice extent curves, the sources of this trend, other than at their annual extrema? We surmised that the trends of the individual annual ice extents could only be discerned by analysing all the data on the ice extents. Perhaps the answer lay in the shape of the individual annual ice-extent oscillations.

Here we present the entire set of these SMMR observations of Arctic and Antarctic sea-ice extent and enclosed open water, which have been corrected for instrumental drift and variations with ecliptic angle. We also give the results of an analysis of the full data set based on band-limited regression⁵ (BLR) which succeeds in accurately determining their trends allowing for the serial correlation present in the raw data.

Instrument stability

Since the earlier publication⁴, there has been additional analysis of the SMMR instrument drift (P.G. *et al.*, manuscript in preparation). First, the annual and semi-annual cycles were removed from the measurements of average radiance from the oceans from 50° N and 50° S to the edges of the Arctic and Antarctic ice packs, respectively. Second, the orbital data for the ascending and descending nodes were averaged separately in each hemisphere to observe the seasonal variations of solar effects on the instrument⁶. This procedure for removing the annual and semi-annual cycles also removes the nine-year means, and when the four results for the ascending and descending, north and south polar averages were compared, they were found to be nearly identical.

An average of the polar ocean radiances between the 50° parallels and the ice edges was the radiation reference used for the detection and subsequent correction of SMMR instrument drift (P.G. *et al.*, manuscript in preparation) in the four channels used in the calculations of sea-ice concentration³: the horizontally and vertically polarized components at wavelengths of 0.8 and 1.7 cm were used. Briefly, the instrument drift was determined on the premise that the averaged polar ocean radiances are invariant with time, and so the observed drift in these ocean radiances constitutes instrument drift. The relative changes observed in each of the channels support this premise. The

TABLE 1 Band-limited regression analysis of polar sea-ice covers

	Slope ($10^6 \text{ km}^2 \text{ yr}^{-1}$)	Standard deviation ($10^6 \text{ km}^2 \text{ yr}^{-1}$)	8.8-yr trend	Confidence level
Arctic sea-ice extent	–0.0315	0.0143	$-2.1 \pm 0.9\%$	96.5%
Antarctic sea-ice extent	+0.0177	0.0188	*	*
Arctic leads and polynyas	–0.0102	0.0058	$-3.5 \pm 2.0\%$	93.5%
Antarctic leads and polynyas	+0.0020	0.0044	*	*

* Statistically insignificant.

premise does not preclude observing changes in the ice pack because the increases in sea surface temperature needed to melt the pack ice are so small that they may not be observable in the SMMR record. As the radiometers are of the Dicke type, no drifts are expected when observing targets radiometrically equivalent to the internal warm references, and none were found. The actual correction for drift was then scaled according to where the observed radiance fell between the values for the ocean and warm reference.

Sea-ice extents and open-water areas

The SMMR radiances used in this analysis have been corrected for the aforementioned drifts and for variations with ecliptic angle⁶. In addition, we have modified the earlier ice-extent calculations by using the 15% ice concentration isopleth as the ice edge and employed a weather filter⁷ to reduce false ice signatures over the oceans due to storms. We decided to use the 15% isopleth for the ice edge to be consistent with similar analyses of the Nimbus-5 microwave data^{8,9}. In the earlier SMMR study⁴, only the weather filter was used, which places the ice edge between the 8% and 12% ice concentration isopleths, depending on ice type. Finally, we improved on the land masks used earlier⁴ for the SMMR data, and we also improved the ice extent calculations by using both the equatorial and polar Earth radii rather than just the equatorial radius as was done in the Nimbus-5 studies. The net result of these improvements is that the SMMR ice-extent variations shown in Fig. 1 are ~7% greater in the Arctic than those reported earlier⁴, and about the same

in the Antarctic. With proper use of the Earth's radii, the Nimbus-5 data for the Arctic ice extent would have been ~1% greater.

The ice-extent and open-water variations in the Arctic and Antarctic were determined with all of the corrected SMMR data obtained over the polar ice covers. On average, data were taken every other calendar day, but there were brief intervals of every-day operation or no operation, occurring in less than 1% of the data set. An uneven data set is not suitable for the analysis used here, and so the earlier data set was edited to produce an evenly spaced time series of alternate days by removal of extra days (1% of the original data) and interpolation to fill the few gaps, adding 1% to the original data.

Trends in the ice covers

A band-limited regression (BLR) technique⁵ has been used previously to establish coherence between atmospheric carbon dioxide and global temperature¹⁰. We will use this technique to obtain an offset and a trend in the SMMR data, which are serially correlated and therefore cannot be analysed properly by standard linear regression. Briefly, the method entails smoothing the data with a multiband filter, and then determining the trend in the filtered data, from which the high-frequency fluctuations and seasonal cycle have been removed. The basis for this technique is the matrix, U .

$$U_{ij} = [\sin 2\pi W(i-j)/\pi(i-j)]; \quad i, j = 0, 1, 2, \dots, N-1 \quad (1)$$

where $0 < W \leq 1/2$ is a bandwidth parameter, and $N = 1,612$ is

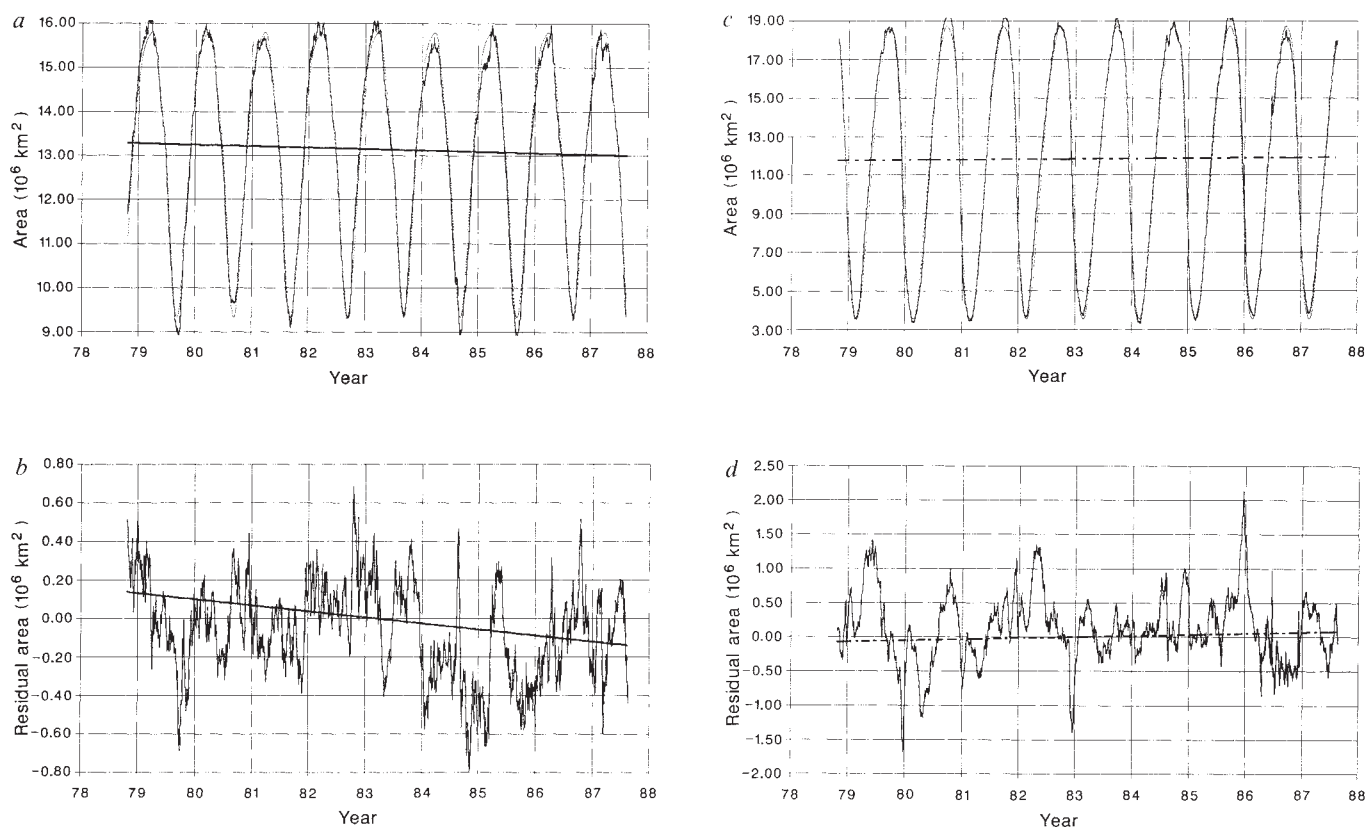


FIG. 1 Sea-ice extents (the areas enclosed by the margin of the sea-ice covers) and residuals for *a* and *b*, the Arctic and, *c* and *d*, the Antarctic from 25 October 1978 to 20 August 1987, and their trends. The Arctic trend (solid line) is a 2.1% decline over the 8.8-year interval under discussion, and is significant to the 96% confidence level. The Antarctic trend (dashed line) is statistically insignificant. The extents are determined from the horizontally and vertically polarized radiances at wavelengths of 0.8 and

1.7 cm obtained by the SMMR on board the Nimbus-7 satellite. Also shown in *a* and *c* are the modelled seasonal oscillations composed of ten sinusoids comprising the first five harmonics of the annual cycle, whose amplitudes were obtained by an ordinary least-squares fit to the data. The residuals shown in *b* and *d* result from subtracting these modelled oscillations from the data, and emphasize the superimposed BLR trend lines. The radiances are corrected for long-term instrument drift and variations with ecliptic angle.

the number of observations. W was chosen so that $W = 4/N$. The bandpass filter used in this approach consists of the first eight eigenvectors and the corresponding eigenvalues of the matrix, U , obtained from its singular value decomposition¹¹, designated as an $8 \times 1,612$ matrix, V , and an 8×8 diagonal matrix, S , respectively. Because of the rapid falloff of the magnitudes of the eigenvalues, a very close approximation (called elsewhere truncated series regression⁵) to the original matrix, U , is obtained as follows:

$$U = VSV' \quad (2)$$

where the prime indicates a matrix transpose. The method continues by formulating the offset and trend in a $1,612 \times 2$ design matrix, D , and obtaining the offset and trend in a coefficient vector, $\mathbf{a}$, as follows:

$$\mathbf{a} = (D'UD)^{-1}(D'U\mathbf{y}) \quad (3)$$

where $\mathbf{y}$ is the data vector. The standard deviations of $\mathbf{a}$ are also obtained from D , V , and $\mathbf{y}$, in a manner⁵ too lengthy to be outlined here.

When this method is applied to the data on the extent of Arctic sea ice obtained from SMMR, a trend of $-0.0315 \times 10^6 \text{ km}^2$ per year is found, with a standard deviation of $0.0143 \times 10^6 \text{ km}^2$ per year. Over the SMMR operational period of 8.8 years, this represents a $2.1 \pm 0.9\%$ decline in the Arctic sea-ice extent. As the basis function of eight eigenvalues was used to determine an offset and trend, the net number of degrees of freedom to be used in the statistical, single-sided t -test⁵ is six, and the confidence level of the determination is 96.5%, based on the ratio of the trend to the standard deviation, which is 2.20. The trend so determined is smaller than the -4% and -5% trends reported earlier⁴ for the maximum and minima, respectively, of the Arctic sea-ice extent, but as mentioned above, these earlier calculations were based on a standard linear regression analysis of uncorrected SMMR data of annual extrema and a lower (8–12%) ice concentration for the ice edge than the 15% used here. The Arctic decline reported here is greater than the -1.4% reported recently¹² for an eight-year interval of annually averaged uncorrected SMMR data. On the other hand, we have found that annually averaging the corrected data over an interval of eight years (so as to average over complete seasonal cycles) before running a linear regression, as done with the uncorrected data earlier¹², gives a trend of -2.2% for the Arctic extent, but the statistical significance of the trend so determined is not valid.

Application of the BLR technique to Arctic open-water area within the ice cover yields a trend of $-3.5 \pm 2.0\%$, at a confidence level of 93.5%. Thus the trends for both the Arctic ice extent and the open water are negative. These trends indicate that for the Arctic the area covered by sea ice decreased less than its ice extent. Application of the technique to the Antarctic sea-ice extent and open-water area leads to statistically insignificant results, as the standard deviations of the trends are greater than the trends themselves. The results are summarized in Table 1 and illustrated in Figs 1 and 2.

We are confident that the BLR analysis reported here more accurately represents the actual trends than results based on ordinary linear regression analysis of uncorrected SMMR data, whether on the annual averages or annual extrema.

Discussion

A number of studies have indicated that changes in the global average air temperature might be detectable by observing changes in the extents of the polar sea-ice covers^{8,9,13–17,23}. The Nimbus-7 SMMR observations give us a unique look at the variations of these extents every two days for almost a decade, and thus provide the most accurate and comprehensive such record yet obtained. The trend obtained by BLR analysis of the Arctic sea-ice extent (Table 1) may be a signal of climate change. It should be said that the near-decade of SMMR observations is not sufficiently long to establish a climate trend clearly. What

sea-ice data are available to extend the record? The interval 1973–76 was covered by an earlier passive microwave radiometer flown in space, the electrically scanned microwave radiometer (ESMR) on board the Nimbus-5 satellite. We have not attempted to treat the ESMR data in the same manner as the SMMR data, because they are not as accurate. The sea-ice record for the time before ESMR is made up largely of data from visible-light sensors flown on the early meteorological satellites; the combination of polar clouds and poor lighting conditions has resulted in a fragmented data set of limited quality for the sea-ice extent.

Whether or not the SMMR record of the decrease in the extents of the Arctic sea-ice cap signals a climate change is therefore an open question. These findings are, however, consistent with records of significant changes in other large-scale phenomena which have occurred over longer periods, including the SMMR decade. A marked increase in the global atmospheric air temperature^{18,19} commenced in 1974 and persisted until the 1980s, which became distinctly the warmest interval in the record lasting for more than 100 years, albeit in the lower latitudes. A warming of the Alaskan Arctic permafrost, typically $2\text{--}4^\circ\text{C}$, commenced during the past 100 years, and during the SMMR lifetime the average increase²⁰ was 0.5°C . A decrease in the extent of the Arctic sea-ice cover has been cited as a possible cause for this warming²⁰, but any permafrost warming corresponding to a change in sea-ice extent would have occurred over much longer times than addressed here. On the other hand, there has been no evidence of warming in ocean surface temperatures in the northern hemisphere during this interval²¹.

We do not know the causes of this asymmetry between the hemispheres, with a significant decrease in Arctic sea-ice extent,

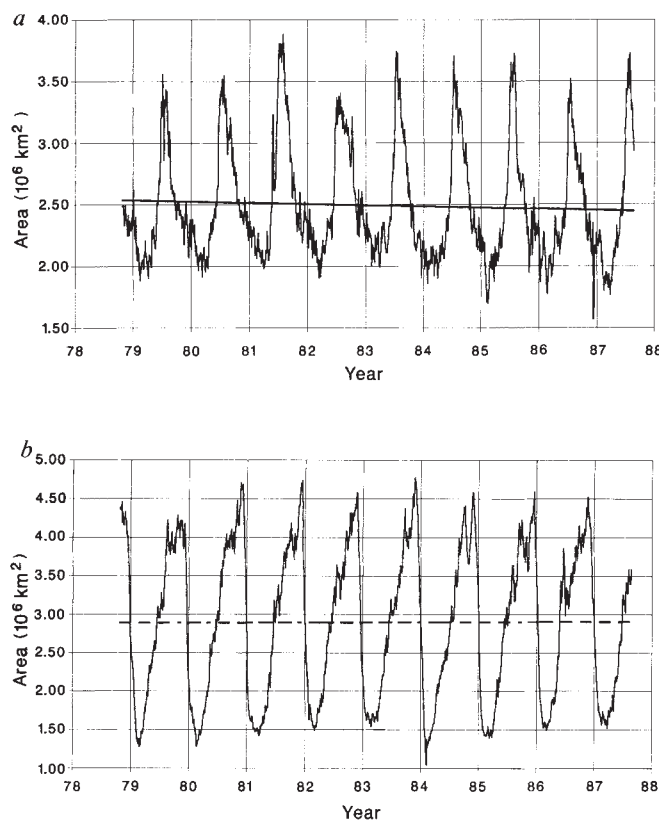


FIG. 2 Open water within the ocean-ice boundary (leads and polynyas) for *a*, the Arctic and *b*, the Antarctic from 25 October 1978 to 20 August 1987, and their trends. The Arctic trend (solid line) is a 3.5% decline over the 8.8-year interval under discussion, and is significant to the 93% confidence level. The Antarctic trend (dashed line) is statistically insignificant.

but no corresponding decrease in the Antarctic. It is interesting to note that a recent simulation with a coupled ocean-atmosphere model²² of the atmospheric response to a doubling of the global CO₂ level over a century shows an interhemispheric asymmetry in the warming of surface air temperatures, with the Antarctic increasing very slowly and the Arctic faster.

We do know that sea ice exists in those parts of the Earth

where global warming is expected to be the greatest. We also know that the sea-ice covers, with their high albedos and insulating effect between the polar atmospheres and oceans, are a key part of the climate machine. That the variations of the Arctic and Antarctic sea-ice extents and enclosed open water have behaved asymmetrically over nearly a decade is a new finding which must be considered in predictions of global change. □

Received 8 April; accepted 31 May 1991.

1. Badgley, F. I. in *Proc. Symp. Arctic Heat Budget and Atmospheric Circulation* (ed. Fletcher, J. O.) 267–277 (1966).
2. Campbell, W. J. *et al. J. geophys. Res.* **92**, 6805–6824 (1987).
3. Cavalieri, D. J., Gloersen, P. & Campbell, W. J. *J. geophys. Res.* **89**, 5355–5369 (1984).
4. Gloersen, P. & Campbell, W. J. *J. geophys. Res.* **93**, 10666–10674 (1988).
5. Lindberg, C. R. *J. R. stat. Soc. B* (submitted).
6. Francis, E. A., thesis, Univ. of Oregon State (1987).
7. Gloersen, P. & Cavalieri, D. J. *J. geophys. Res.* **91**, 3913–3919 (1986).
8. Zwally, H. J. *et al. NASA SP-459* (U.S. GPO, Washington DC, 1983).
9. Parkinson, C. L. *et al. NASA SP-489* (U.S. GPO, Washington DC, 1987).
10. Kuo, C., Lindberg, C. & Thomson, D. J. *Nature* **343**, 709–714, (1990).
11. Press, W. H., Flannery, B. P., Teukolsky, S. A. & Vetterling, W. T. *Numerical Recipes in C* (Cambridge University Press, 1989).
12. Parkinson, C. L. & Cavalieri, D. J. *J. geophys. Res.* **94**, 14499–14523 (1989).

13. Sissla, J. F., Sabatini, R. R. & Ackerman, H. J. *Polar Res.* **16**, 367–373 (1972).
14. Streten, N. A. *Arch. Meteorol. geophys. Bioklimatol. Ser. A22*, 129–134 (1973).
15. Budd, W. F. *J. Glaciol.* **15**, 417–426 (1975).
16. Kukla, G. & Gavin, J. *Science* **215**, 497–503 (1981).
17. Carsey, F. D. *J. geophys. Res.* **87**, 5809–5835 (1982).
18. Jones, P. D. *et al. Nature* **322**, 790 (1988).
19. Hansen, J. & Lebedeff, S. *Geophys. Res. Lett.* **15**, 323–326 (1988).
20. Lachenbruch, A. H., Cladouhos, T. T. & Saltus, R. W. *Permafrost* **3**, 9–17 (ed. Senneset, K.) (Tapir, Trondheim, Norway, 1988).
21. Reynolds, W. R., Folland, C. K. & Parker, D. E. *Nature* **341**, 728–731 (1989).
22. Stouffer, R. J., Manabe, S. & Bryan, K. *Nature* **342**, 660–662 (1989).
23. Zwally, H. J., Parkinson, C. L. & Comiso, J. C. *Science* **220**, 1005–1012 (1983).

ACKNOWLEDGEMENTS. We thank C. R. Lindberg of AT&T Bell Laboratories, H. F. Lins of the US Geological Survey Water Resources Division and A. J. Busalacchi of the NASA Goddard Laboratory for Hydrospheric Processes for critical readings of this paper and helpful suggestions.

Chaperonin-mediated protein folding at the surface of groEL through a 'molten globule'-like intermediate

Jörg Martin, Thomas Langer, Raina Boteva*, Andrea Schramel*, Arthur L. Horwich† & F.-Ulrich Hartl‡

Institut für Physiologische Chemie der Universität München, Goethestrasse 33, W-8000 München 2, Germany

* Zoologisches Institut der Universität München, Luisenstrasse 14, W-8000 München 2, Germany

† Howard Hughes Medical Institute and Department of Human Genetics, Yale School of Medicine, 333 Cedar Street, New Haven, Connecticut 06510, USA

Folding of two monomeric enzymes mediated by groE has been reconstituted *in vitro*. The groEL protein stabilizes the polypeptides in a conformation resembling the 'molten globule' state. Mg-ATP and groES then promote the acquisition of ordered tertiary structure at the surface of groEL. Folding requires the hydrolysis of about 100 ATP molecules per protein monomer. This active process of surface-mediated chain folding might represent a general mechanism for the formation of protein structure *in vivo*.

LITTLE is known about the mechanisms by which newly synthesized proteins fold inside cells. For many proteins, folding *in vivo* may not be a spontaneous process. Several components are involved in mediating protein folding in a variety of cell types and compartments^{1–3}. With the exception of protein disulphide isomerase and peptidyl prolyl isomerase, they have been classified as 'molecular chaperones'^{3,4} or 'polypeptide chain binding proteins'⁵ on the basis of their ability to prevent the formation of wrong aggregates by binding to unfolded or partially denatured proteins. The heat-shock proteins of the hsp70

and hsp60 families are typical representatives of this group of components^{5–7}.

The essential function in protein folding of the members of the hsp60 family has been demonstrated *in vivo* and *in vitro*. These so-called 'chaperonins'⁸ comprise the groEL protein of *Escherichia coli* and other bacteria^{8–17}, the rubisco-binding protein of chloroplasts^{18–20} and the mitochondrial hsp60^{21–25}. They form complexes with ATPase activity composed of 14 subunits with a relative molecular mass of 60,000 (*M_r* 60K) which are arranged in two stacked seven-subunit rings. The groEL protein and the mitochondrial hsp60 functionally cooperate with groES^{26,27}, a ring-shaped complex of seven roughly 10K subunits, which inhibits the ATPase activity of groEL^{26,28}. The groE proteins are required for λ-phage head assembly^{9,10}. The ribulose 1,5-bisphosphate carboxylase oxygenase (Rubisco) binding-protein mediates the assembly of hexadecameric Rubisco in chloroplasts^{18–20}. The groE-dependent assembly of dimeric prokaryotic Rubisco has been reconstituted *in vitro*^{13,28}. The mitochondrial hsp60 is necessary not only for the oligomeric assembly of proteins^{23,25} but also for the chain folding of monomeric polypeptides²⁴. The molecular mechanism of this ATP-driven process remains unknown.

Using the groEL and groES proteins of *E. coli*, we have reproduced the chaperonin-dependent folding of two monomeric enzymes, dihydrofolate reductase (DHFR) and rhodanese. Folding reactions were monitored by measuring intrinsic tryptophan fluorescence, adsorption of the fluorescent dye anilino-naphthalene-sulphonate, protease sensitivity and enzyme activity. We find that groEL stabilizes an early intermediate on the folding pathway seemingly equivalent to the folding

‡ To whom correspondence should be addressed. Address after November 1: Laboratory of Cellular Biochemistry, Rockefeller Research Laboratory, Sloan-Kettering Institute, New York, New York 10021, USA.

state described as 'molten globule'²⁹⁻³²; that ATP-dependent folding occurs at the surface of groEL through intermediate conformations which are progressively more compact but still enzymatically inactive; and that by regulating the groEL ATPase, groES accomplishes a critical folding steps at groEL. We propose that the mechanism of chaperonin-catalysed folding involves stepwise ATP-dependent release of the protein substrate from the groEL scaffold.

Reconstitution of chaperonin-dependent folding

Chicken DHFR is a globular, single-domain protein of 189 residues^{33,34}. On dilution from denaturant, more than 90% of total DHFR refolds³⁵ (Fig. 1). We used this property to establish the conditions allowing binding of DHFR to purified groEL. All DHFR was unfolded in 6 M guanidinium-HCl (GdmCl), as indicated by circular dichroism spectroscopy (not shown), and it was then diluted 200-fold into groEL-containing buffer (Fig. 1a). An equimolar concentration of groEL 14-mer complex reduced the yield of reactivated DHFR by about 80%. Unfolded DHFR bound to groEL with high affinity independent of the presence of ATP (in the absence of Mg²⁺). As for other denatured proteins^{15,16}, the groEL 14-mer binds only one or two

molecules of DHFR (Fig. 1a, insert). Addition of Mg-ATP to the DHFR-groEL complex allowed refolding to occur with a half-time ($t_{1/2}$) of about 3 min (Fig. 1b). This was considerably slower than the spontaneous folding of DHFR ($t_{1/2}$ of about 1 min). Curves for reactivation in the presence of groEL were sigmoidal, suggesting the existence of additional species of DHFR other than the fully unfolded and native forms. The groES 7-mer added at a 1:1 molar ratio to groEL accelerated the groEL-dependent folding. Higher amounts of groES had no further effect.

Increasing the concentration of groEL in the assay retarded the Mg-ATP-dependent reactivation of DHFR (Fig. 1c). This suggested that, after release, DHFR could re-bind to groEL before folding is complete. We tested this using α_{S1} -casein as a competitor of protein binding to groEL (Fig. 1d). Casein has certain properties of a partially denatured protein. Although being soluble, the native form exposes a considerable part of its hydrophobic residues to solvent and contains a high amount of disordered structure^{36,37}. Casein bound to groEL with high affinity and thereby prevented the rebinding of DHFR to groEL. Displacement of DHFR by casein was only detected in the presence of Mg-ATP. Notably, the decreased rate of reactivation

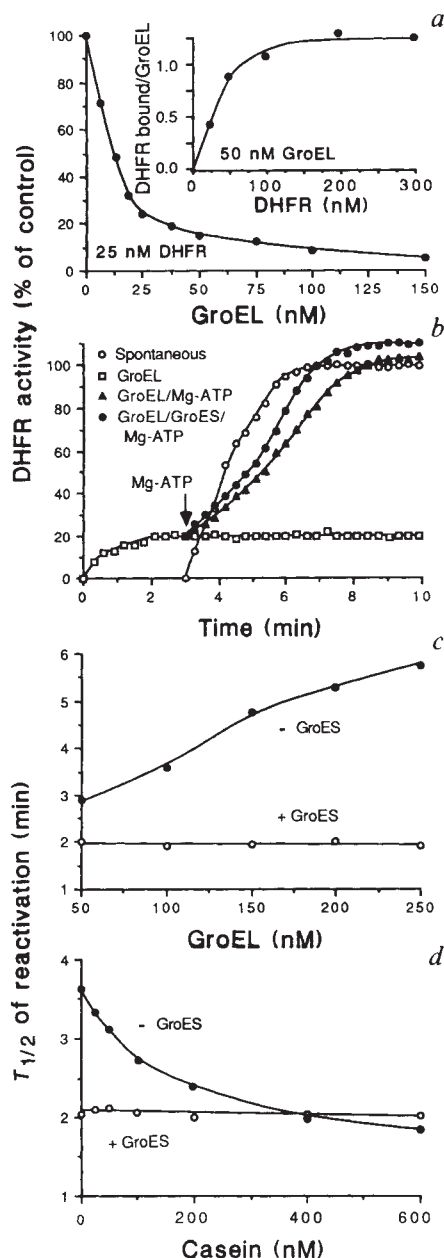


FIG. 1 GroE-dependent reactivation of unfolded DHFR. *a*, Inhibition of spontaneous refolding by groEL. Insert, binding of DHFR to groEL. *b*, Time course of reactivation in the absence (○) or presence of groEL (□, ▲, ●), groES (●) and Mg-ATP (▲, ●). Influence of increasing groEL concentrations (*c*) and effect of casein (*d*) on $t_{1/2}$ of Mg-ATP-dependent reactivation with (○) and without (●) groES. The activity of DHFR on spontaneous refolding is set to 100%.

METHODS. The groEL and groES proteins were purified to >95% purity^{15,28} using a groE-overproducing strain of *Escherichia coli* harbouring the plasmid pOF39 (ref. 50). *a*, Chicken DHFR (>97% purity, Sigma) was denatured at 100 $\mu\text{g ml}^{-1}$ in buffer A (6 M guanidinium-HCl (GdmCl), 30 mM Tris buffer pH 7.4, 1 mM dithiothreitol (DTT)) and diluted 200-fold at 15 °C into buffer B (30 mM Tris, pH 7.2, 2 mM DTT, 50 mM KCl) containing 50 μM NADPH, 50 μM dihydrofolate and groEL 14-mer as indicated. DHFR activities were determined photometrically by monitoring the decrease in absorbance at 340 nm (ref. 42). Activities indicated were measured 3 min after dilution when refolding of free DHFR was complete. Binding to groEL was analysed by adding denatured DHFR at 15 °C to buffer B containing 50 nM groEL (GdmCl, 100 mM final concentration). After 5-min incubation, reactions were separated on 5 ml Sephacryl S300-columns (Pharmacia)²⁴. Fractions of 120 μl were analysed by SDS-PAGE and immunoblotting using anti-DHFR and anti-groEL antisera, and the luminescence-based detection system ECL (Amersham)⁵¹. Fluorographs were quantified by laser-densitometry. *b*, Unfolded DHFR was diluted to 25 nM into buffer B with or without 50 nM groEL. Reactivation was started after 3 min (arrow) by adding 5 mM magnesium acetate and 1 mM ATP. When indicated, groES 7-mer was added to 50 nM before Mg-ATP. Progress curves of activity were recorded by plotting ΔA_{340} per 17 s versus time. *c*, The DHFR-groEL complex was generated as in (*a*) (groEL/DHFR, 50 nM/25 nM). After 3-min incubation, increasing amounts of groEL were added in the presence or absence of equimolar groES (total concentrations of groEL indicated). Reactivation was initiated by adding Mg-ATP. On the basis of progress curves, $t_{1/2}$ is the time required for 60% reactivation (spontaneous reactivation set to 100%) considering that ~20% of total DHFR had escaped binding to groEL. *d*, Unfolded DHFR (25 nM) was diluted into buffer B containing 100 nM groEL in the presence or absence of groES (100 nM). After 3 min, 0–600 mM α_{S1} -casein (Sigma) was added. After further incubation for 3 min, Mg-ATP was added and $t_{1/2}$ of reactivation determined as in *c*.

due to the rebinding of DHFR was not observed when groES was present in amounts stoichiometric to groEL (Fig. 1c, d). It seems that only when groES is absent does DHFR reach its native state through many cycles of binding to groEL and Mg-ATP-dependent release only when groES is absent.

In contrast to DHFR, many proteins do not refold efficiently on dilution from denaturant; rather, they form misfolded aggregates^{17,32,38}. The groEL and groES proteins mediated the

refolding of one such monomeric enzyme, rhodanese of bovine liver. Rhodanese (293 residues) is composed of two equal-sized domains which are stabilized by hydrophobic interactions³⁹. Dilution of rhodanese from 6 M GdmCl into buffer alone resulted in aggregation⁴⁰ (Fig. 2a). GroEL prevented this aggregation by complex formation with rhodanese. Unlike DHFR, only a small amount of rhodanese was released from groEL on ATP hydrolysis; this portion of rhodanese did not reactivate but formed inactive aggregates (Fig. 2b). Complete ATP-dependent release of rhodanese was achieved by the competitor casein and resulted in the aggregation of the released protein. Only in the presence of groES did efficient reactivation of rhodanese ($t_{1/2} \sim 10$ min) occur (Fig. 2c), suggesting a specific function of groES for folding beyond just promoting the release of a groEL-stabilized intermediate.

GroEL stabilizing a 'molten globule'-like state

What is the conformation of the polypeptide in association with groEL? We examined the intrinsic tryptophan fluorescence of groEL-bound DHFR and rhodanese as a measure of tertiary structure. The groEL and groES components interfere only minimally with this analysis because both lack Trp residues^{8,41}.

DHFR from chicken contains tryptophans at positions 24, 57 and 113 (ref. 33). Their fluorescence changes markedly when the protein is unfolded: the emission maximum ($\lambda_{\max,em}$) shifts from 331 nm to 356 nm accompanied by a roughly twofold increase in the fluorescence intensity (Fig. 3a). The redshift of fluorescence generally reflects transfer of Trp residues to a more polar environment. The increase in fluorescence intensity on unfolding is largely due to the removal of quenching amino acids from near Trp 24 (ref. 42). The $\lambda_{\max,em}$ of DHFR bound to groEL in the absence of ATP hydrolysis (termed DHFR-11) was 345 nm (Fig. 3a), halfway between the fluorescence maxima of the native and unfolded proteins (Table 1). Strikingly, the relative fluorescence intensity of DHFR-11 was almost as high as that of the completely unfolded protein. Apparently, groEL-bound DHFR lacks an ordered tertiary structure. This is consistent with the very high protease sensitivity of the protein (Fig. 3c). The Trp residues in DHFR-11 are already in a more hydrophobic environment than the completely unfolded protein, as

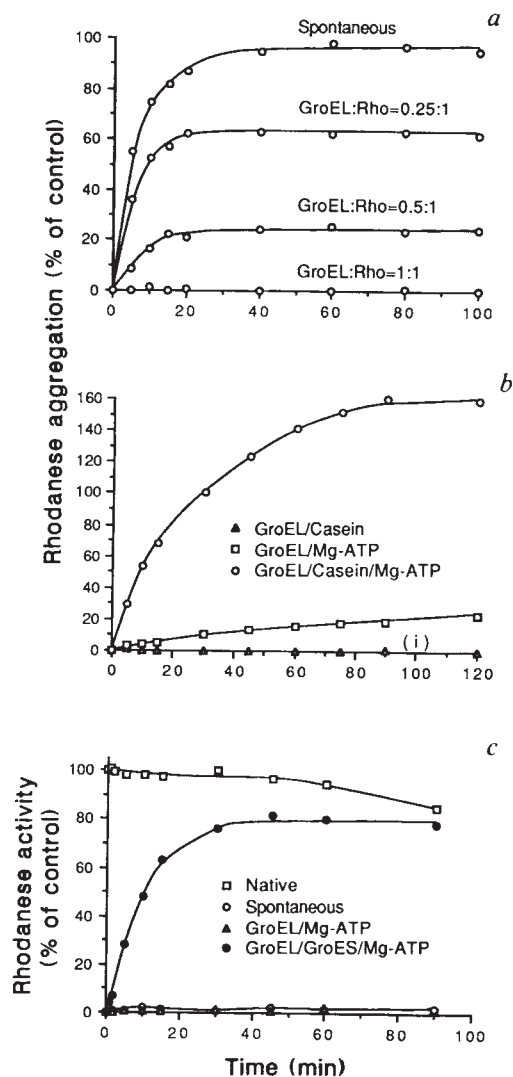


FIG. 2 GroE-dependent reactivation of rhodanese. a, Suppression of aggregation by groEL. b, Displacement of rhodanese from groEL by casein (Δ , $\circ$) in the presence of Mg-ATP ($\square$, $\circ$). Aggregation in the absence of groEL is set to 100%. c, Reactivation of unfolded rhodanese in the absence ($\circ$) and presence of groEL and Mg-ATP (Δ , $\bullet$) and groES ($\bullet$). The activity of native rhodanese ($\square$) is set to 100%.

METHODS. Rhodanese from bovine liver (>95% purity, Sigma) was denatured in buffer A. a, Unfolded rhodanese was diluted 100-fold (0.46 μ M final concentration) into buffer C (30 mM Tris, pH 7.2, 50 mM KCl) containing 0, 0.12, 0.23 and 0.46 μ M groEL 14-mer, respectively. Aggregation (turbidity) was measured as absorbance at 320 nm (ref. 52). b, Rhodanese-groEL complex was generated as above (rhodanese/groEL, 0.46 μ M/0.69 μ M). Aggregation was measured as in (a) on addition of 1 mM ATP and 5 mM magnesium acetate in the absence or presence of 4.4 μ M casein. (Aggregates could be sedimented while groEL remained in the supernatant fraction.) c, Rhodanese-groEL complex was formed as in b. When indicated, groES was present at 0.69 μ M. At various times after adding Mg-ATP, enzyme activities were determined⁴⁴ in the presence of 10 mM *trans*-1,2-cyclohexanediaminetetraacetate (CDTA) to stop ATP-hydrolysis.

TABLE 1 Fluorescence parameters of DHFR and rhodanese species

	Maximum of tryptophan fluorescence (nm)	Fluorescence at emission maximum (arbitrary units)	K_Q (1/M)	f_a
DHFR				
Native	331	6.7	2.5	1.0
6 M GdmCl	356	15.5	8.5	1.0
GroEL-bound (DHFR-11)	345	14.1	5.5	0.8
Rhodanese				
Native	333	13.1	1.5	1.0
6 M GdmCl	356	12.2	9.0	1.0
GroEL-bound (RHO-11)	342	17.8	4.1	1.0

Quenching of Trp fluorescence using 0–400 mM acrylamide was measured at the emission maximum (see Fig. 3). Quenching data were analysed according to the Stern-Volmer relationship $F_0/F = 1 + K_Q[X]$, where F_0 and F are the fluorescence intensities in the absence and presence of the quencher, respectively, and $[X]$ is the quencher concentration. K_Q is the apparent quenching constant. In the case of heterogeneous emission, a modified Stern-Volmer equation was applied to calculate the effective quenching constant ($K_{Q,eff}$) and the fraction of Trp residues (f_a) with the highest accessibility to the quencher⁵⁴. Data were corrected for background fluorescence in the absence of substrate proteins and for the inner-filter effect due to acrylamide.

indicated by the 50% blueshift of emission: but using acrylamide as quencher indicated that their accessibility to solvent was much greater than that of the native structure (Table 1). Acrylamide is a polar but non-ionic collisional quencher that may partition into the loosely-folded interior of folding intermediates.

The properties of groEL-bound rhodanese (RHO-I1) were surprisingly similar to those of DHFR-I1. Rhodanese contains eight tryptophan residues³⁹: on complete unfolding, their fluorescence maximum shifts from 330 nm to 355 nm without major change in intensity (Fig. 3a). The $\lambda_{\text{max,em}}$ of RHO-I1 was

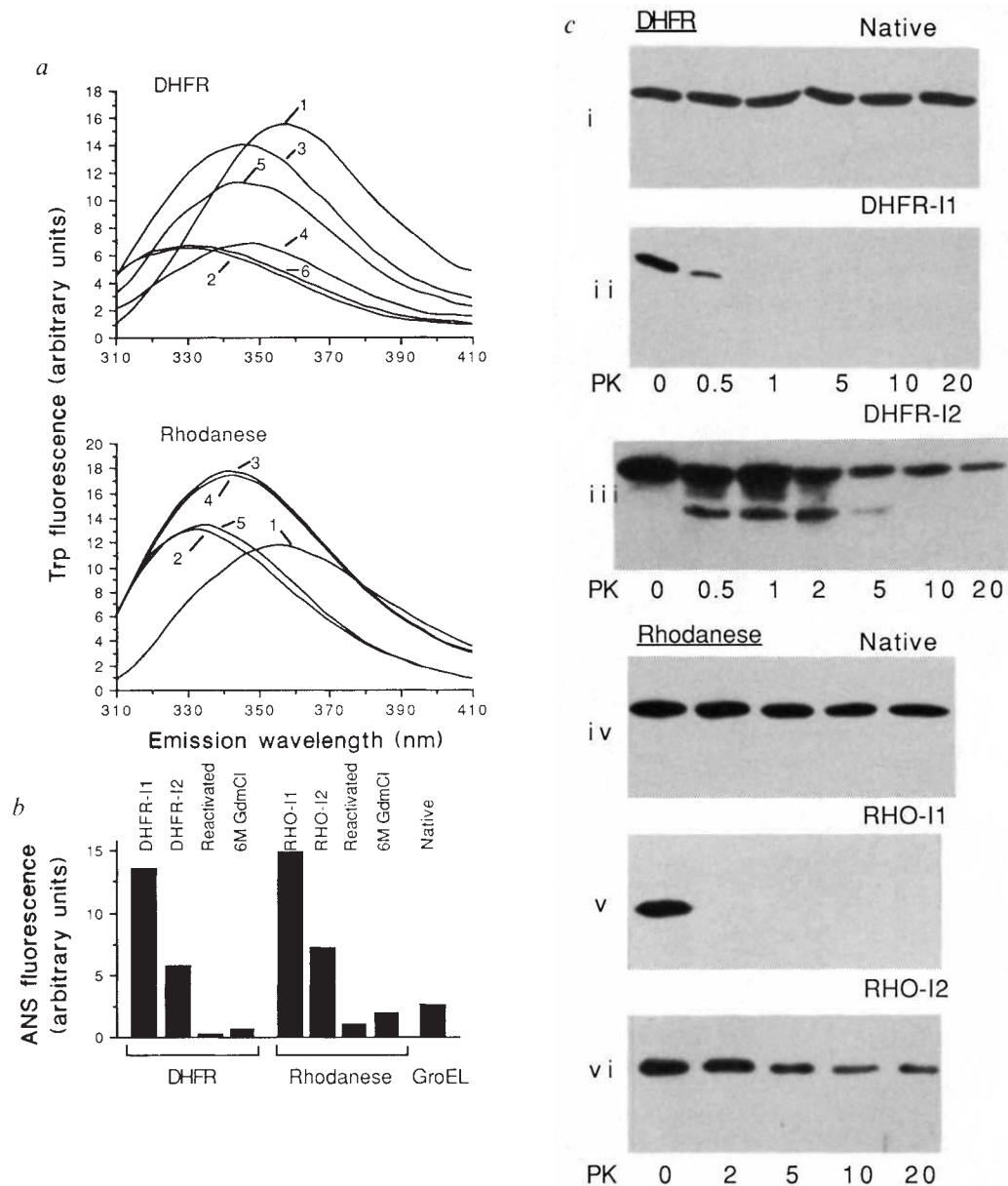
at 342 nm, again about 50% shifted towards that of the unfolded form. The fluorescence intensity of RHO-I1 was increased and was characterized by the presence of partially accessible Trp residues (Table 1). Like DHFR-I1, groEL-bound rhodanese was very sensitive to protease (Fig. 3c).

We tested whether DHFR-I1 and RHO-I1 contained hydrophobic regions that could be detected by 1-anilino-naphthalene-8-sulphonate (ANS), a probe for apolar binding sites whose fluorescence is strongly dependent on the hydrophobicity of the environment. ANS accumulates in the solvated hydrophobic core of early folding intermediates generally termed 'molten

FIG. 3 Conformations of groEL-associated proteins. **a**, Upper plot: Trp fluorescence of DHFR unfolded in 6 M GdmCl (1); DHFR native (2); groEL-bound DHFR-I1 (3); groEL-bound DHFR-I2 (4); DHFR-I2 after incubation in the presence of CDTA (5); DHFR after complete reactivation with groEL and Mg-ATP (6). Bottom plot: Trp fluorescence of rhodanese unfolded in 6 M GdmCl (1); native rhodanese (2); groEL-bound RHO-I1 (3); RHO-I1 on incubation with Mg-ATP in the absence of groES (4); RHO-I1 after reactivation in the presence of groES and Mg-ATP (5). **b**, ANS (1-anilino-naphthalene-8-sulphonate) fluorescence of free and groEL-bound proteins. **c**, Proteinase K resistance of native DHFR and native rhodanese (i, iv); DHFR-I1 and RHO-I1 (ii, v); DHFR-I2 and RHO-I2 (iii, vi) (respective numerals in parentheses).

METHODS. Fluorescence spectra were recorded on a SPEX Fluorolog 211. Absorbance was lower than 0.05 at the excitation wavelength. Protein-groEL complexes were formed as above at a 2.5-fold molar excess of groEL 14-mer over substrate protein (DHFR and rhodanese 0.65 μM and 0.3 μM final concentration, respectively). More than 90% of DHFR and 95% of rhodanese were groEL-bound. Where indicated, groES 7-mer was added in equimolar concentrations to groEL. **a**, Trp fluorescence was excited at 295 nm. Background fluorescence of chemically identical reactions lacking substrate proteins (15–30% of total fluorescence due to minor impurities of the groE preparation) was subtracted. DHFR which had escaped binding to groEL was corrected for. Analysis was at 15 °C (DHFR) and 20 °C (rhodanese). For **a**, top plot, DHFR-I2 in reaction 4 was generated by adding Mg-ATP (1 mM

ATP/2 mM magnesium acetate) to DHFR-I1. Reactions 4 and 6 were recorded 30 s and 10 min after adding Mg-ATP, respectively. Unfolding of DHFR-I2 in reaction 5 was observed after 5-min incubation of reaction 4 with 10 mM CDTA. For **a**, bottom trace, RHO-I1 in reaction 4 was incubated with Mg-ATP for 30 min. Reactivated rhodanese in reaction 5 was analysed after 40-min incubation of RHO-I1 with groES and Mg-ATP. Various forms of DHFR and rhodanese were incubated for 5 min at 15 °C or 20 °C, respectively, with a 20-fold molar excess of ANS^{29,43}. Emission spectra (excitation, 390 nm; emission, 470 nm) were corrected for background fluorescence caused by ANS in reactions lacking DHFR or rhodanese. DHFR-I2 and reactivated DHFR were formed as above in the presence of ANS and fluorescence at 470 nm was followed. RHO-I2 was generated in the presence of ANS by incubation



of RHO-I1 with Mg-ATP and groES for 3 min at 20 °C. Fluorescence intensity is given in arbitrary units per mole of protein (for 14-mer in case of groEL). **c**, GroEL-bound DHFR-I1 and DHFR-I2 were separated from free DHFR by sizing chromatography as in Fig. 1. GroEL-bound rhodanese was isolated by sedimenting aggregates for 30 min at 25,000g. RHO-I2 was generated as above (but in the absence of ANS). The reaction was stopped by adding 10 mM CDTA and cooling to 0 °C. Protease treatment was performed in buffer B (DHFR and rhodanese, ~0.25 μM) for 10 min at 0 °C at 0–20 $\mu\text{g ml}^{-1}$ proteinase K (PK). Protease was inhibited by 2 mM phenylmethylsulfonyl fluoride. TCA precipitates were analysed by SDS-PAGE and immunoblotting (see Fig. 1). The fluorograph in reaction iii was exposed for 90 s instead of 15 s.

globule^{29,43}. Both groEL-stabilized proteins showed strong ANS fluorescence (Fig. 3b). The stability of the complex between groEL and the substrate proteins was preserved in the presence of ANS. There was little fluorescence with the native or guanidine-unfolded proteins, or with the groEL 14-mer alone. This does not exclude the possibility that groEL shields hydrophobic regions of the associated proteins, thereby increasing the hydrophobic space accessible to ANS. The near spatial relationship of the tryptophan and ANS chromophores in DHFR-I1 and RHO-I1 is indicated by the efficiency of fluorescence energy transfer from the Trp residues to ANS being about 30% (not shown). This again suggests that the two groEL-bound proteins have a similar overall shape.

These results indicate that groEL stabilizes unfolded substrate

proteins in a conformation devoid of ordered tertiary structure. The fluorescence properties and the increased solvent accessibility of the Trp residues, as well as the incorporation of ANS, point to the presence of flexible tertiary structure internalizing a fluctuating, hydrophobic core, as described for the molten globule state.

DHFR folding through a more compact intermediate

We monitored the changes in Trp fluorescence during groEL-dependent folding of DHFR-I1. Within 15–30 s after the addition of Mg-ATP to the DHFR-groEL complex, the relative fluorescence intensity of DHFR-I1 dropped markedly, whereas $\lambda_{\max,em}$ was unaltered, defining the folding intermediate DHFR-I2 (Fig. 3a). In a subsequent period of 5–10 min, $\lambda_{\max,em}$ shifted from 345 nm to 331 nm accompanied by the appearance of enzyme activity. Apparently, a partial folding reaction including Trp 24 occurs early in groEL-dependent folding of DHFR, exposing this residue to the native-like environment.

In parallel with the marked decrease in Trp fluorescence, the conformational change to DHFR-I2 resulted in a loss of 60% of the ANS fluorescence detected with DHFR-I1 (Fig. 3b). This is typically observed with folding intermediates, resembling the molten globule on their organization to a more compact structure^{29,43}. We confirmed this by analysing the endogenous resistance of DHFR-I2 to protease. DHFR-I2 was generated by initiating ATP hydrolysis as above. After 30 s at 15 °C, the folding reaction was stopped by adding the Mg^{2+} chelator CDTA and cooling to 0 °C. DHFR-I2 was resistant to intermediate concentrations of proteinase K (Fig. 3c). A small amount of a ~10K fragment was detectable which was digested at higher concentrations of protease. In contrast to the protease-stable native enzyme, DHFR-I2 was still associated with groEL: when the reaction containing DHFR-I2 and CDTA was maintained at 15 °C (instead of 0 °C), there was partial unfolding, as indicated by an increase in fluorescence intensity (Fig. 3a). Addition of sufficient Mg^{2+} to compensate for the CDTA present reinitiated the folding reaction (not shown). The folding of groEL-bound DHFR-I1 apparently involves the transient formation of an intermediate with partially native tertiary structure, DHFR-I2. This compaction occurs in close association with the groEL complex.

GroES and folding at the groEL surface

The interaction of groES with groEL accelerates the conversion of DHFR-I1 to the native protein (Fig. 1b), perhaps simply by facilitating its Mg-ATP-dependent release. Alternatively, because folding of rhodanese is essentially groES-dependent (Fig. 2c), groES might allow a critical folding step(s) to proceed at the groEL surface. We attempted to dissect rhodanese folding into intermediate steps. Without groES, prolonged incubation with Mg-ATP had no effect on the fluorescence properties (Fig. 3a) or the protease-sensitivity (not shown) of groEL-bound RHO-I1. Folding in the presence of groES was stopped after various times by adding CDTA to inhibit ATP hydrolysis. The amount of groEL-bound rhodanese, its endogenous protease resistance, and enzyme activity were followed (Fig. 4a). During the first 5 min of incubation in the presence of Mg-ATP, an intermediate of rhodanese folding accumulated, defined as RHO-I2, which had lost 50% of the ANS fluorescence of RHO-I1 and had acquired an increased resistance to intermediate concentrations of proteinase K (Fig. 3). The formation of this inactive intermediate preceded that of protease-resistant, active rhodanese which retained only little ANS fluorescence. In contrast to the native enzyme, RHO-I2 was still bound to groEL (Fig. 4a). In summary, RHO-I2 has properties resembling DHFR-I2 but, in contrast to DHFR-I2, its formation is critically dependent on groES.

These experiments did not address whether folding occurred at the surface of groEL without intermittent release from the chaperonin. If folding involved cycles of complete release and

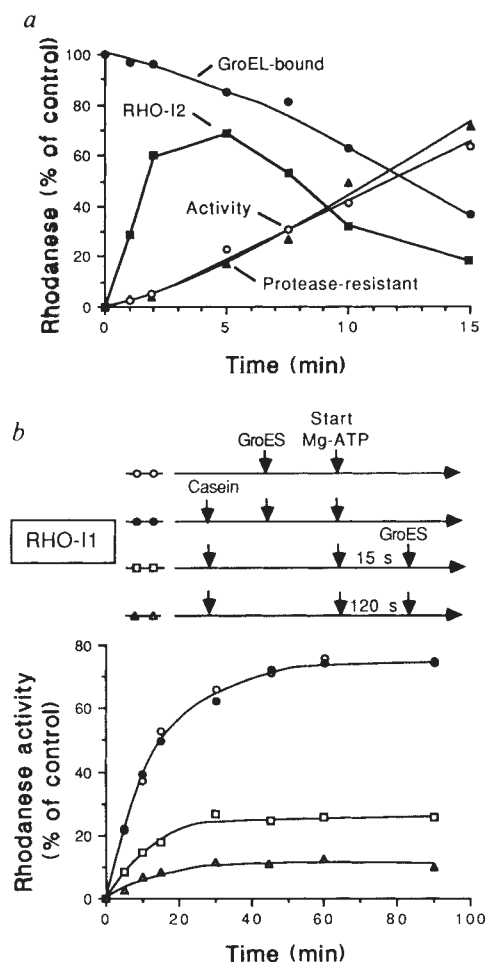


FIG. 4 Folding of rhodanese at the surface of groEL. *a*, Time-course of accumulation of groEL-associated RHO-I2 and of active rhodanese. The total amount of rhodanese analysed is set to 100%. *b*, Reactivation of rhodanese in the absence (○) or presence (●, □, △) of casein adding groES before Mg-ATP (●), and 15 s (□) or 120 s (△) after Mg-ATP.

METHODS. *a*, GdmCl-unfolded rhodanese was diluted to 1.17 μM into buffer C containing 2.34 μM groEL 14-mer. After centrifugation to remove aggregates, 2.34 μM groES was added to the supernatant fraction and refolding was started at 25 °C by adding 1 mM ATP and 2 mM magnesium acetate. At the times indicated, folding was stopped in aliquots of the reaction by adding CDTA and cooling to 0 °C. Enzyme activity was determined and protease resistance was titrated using 0–20 $\mu g ml^{-1}$ proteinase K (see Fig. 3c). RHO-I2 is the amount of rhodanese resistant to 2 $\mu g ml^{-1}$ proteinase K minus the amount of refolded rhodanese resistant to 20 $\mu g ml^{-1}$ proteinase K. Association with groEL was determined as in Fig. 1. *b*, Refolding reactions containing RHO-I1 (0.46 μM rhodanese and 0.69 μM groEL) received 4.4 μM casein, 0.69 μM groES, 1 mM ATP and 5 mM magnesium acetate when indicated.

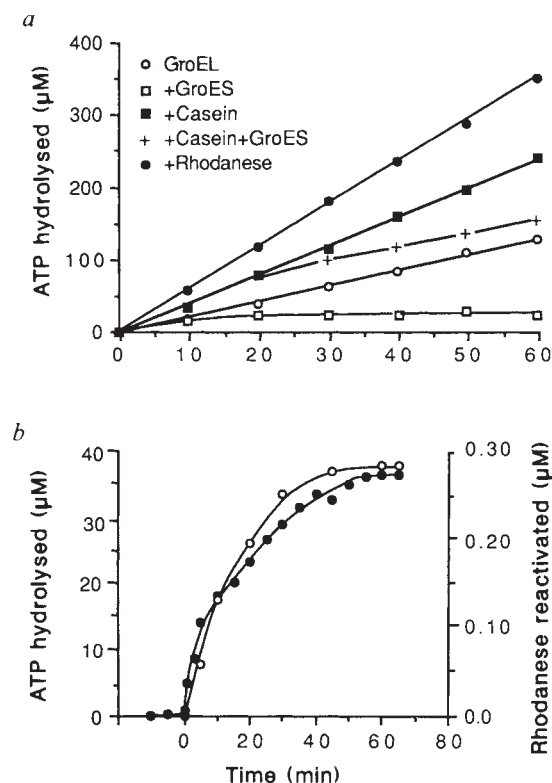


FIG. 5 ATP-requirement of groE-dependent refolding of rhodanese. *a*, Influence of casein, rhodanese and groES on the ATPase activity of groEL: ○, groEL alone; □, with groES; ■, with casein; +, with casein, groES at 20 min; ●, with rhodanese. *b*, Stoichiometry of ATP hydrolysis (●) and groE-dependent reactivation of rhodanese (○).

METHODS. Free phosphate produced by ATP hydrolysis was measured colorimetrically⁵³. Spontaneous hydrolysis of ATP was corrected for. *a*, ATP hydrolysis was started at 25 °C by adding 1 mM ATP and 5 mM magnesium acetate to buffer C containing 0.33 μM groEL, and 0.33 μM groES, 4.4 μM casein or 1.0 μM denatured rhodanese as indicated. GroES was added to the reaction containing casein after 20-min incubation. GdmCl (<30 mM) had no effect on the groEL ATPase. Up to 1.6 μM aggregated rhodanese had no stimulating effect. *b*, GroEL (0.33 μM) and groES (0.66 μM) were incubated with Mg-ATP for 30 min at 25 °C to ensure inhibition of the groEL ATPase. At *t*=0, denatured rhodanese was added to 1.0 μM. Enzyme activity and ATP hydrolysis were measured in aliquots of the same reaction. The concentration of reactivated rhodanese was calculated based on the specific activity of the native enzyme. The results of one out of five experiments are shown.

rebinding to groEL, the addition of casein should compete for the groES-dependent conversion to RHO-12 and to the native enzyme. A six-fold molar excess of casein over groEL efficiently displaced RHO-11 in the presence of Mg-ATP (Fig. 2b). When added before rhodanese, this amount of casein was sufficient to almost completely prevent the reactivation of rhodanese by groEL and groES (not shown). Strikingly, when both casein and groES were added to preformed rhodanese-groEL complex before initiating folding by Mg-ATP, reactivation was as efficient as in the absence of casein (Fig. 4b). Thus, the competitor was unable to interrupt the reactivation once the groEL-rhodanese complex had been formed and groES was present. This suggests that a single round of interaction between rhodanese and groEL is sufficient for folding. But when groES was added to a reaction containing rhodanese-groEL complex and casein 15–120 s after starting ATP hydrolysis, the yield of active enzyme decreased by 50–80% (Fig. 4b). This corresponded to the amount of rhodanese displaced from groEL during that incubation. During folding to RHO-12, rhodanese seems to maintain contact with the groEL 14-mer. In the presence of groES, the protein is committed to fold. The underlying mechanism of this surface-mediated reaction could be a progressive, Mg-ATP-dependent release of the folding polypeptide.

ATP requirement of folding

The ATPase activity of groEL, measurable in the absence of protein substrate, is inhibited by groES^{26,28}. The groEL 14-mer hydrolysed ATP at a rate of about 7 molecules per min (Fig. 5a). The groES 7-mer, at a 1:1 molar ratio to the groEL 14-mer, reduced this rate to about 0.25 molecules per min within 5–10 min after addition. Casein caused a twofold activation of the groEL ATPase in the absence of groES (Fig. 5a). Also, binding of unfolded rhodanese resulted in a continuous stimulation. This is consistent with the finding that, in the absence of groES, rhodanese and groEL undergo cycles of complete Mg-ATP-dependent release and rebinding without reactivation. Under these conditions, the competitor casein will displace 50% of bound rhodanese from groEL within 15 s of ATP hydrolysis (Fig. 4b). We estimate that about 27 molecules of ATP are

hydrolysed per molecule of rhodanese released in the absence of groES.

What is the stoichiometry between ATP hydrolysis and folding at groEL in the presence of groES? Addition of unfolded rhodanese to groEL and groES initially resulted in a roughly 40-fold increase of the groES-suppressed rate of ATP hydrolysis (Fig. 5b). The ATPase activity then gradually decreased as reactivation of rhodanese approached completion. Per mole of rhodanese folded, 130 ± 20 moles of ATP were hydrolysed (Fig. 5b). By contrast, casein caused a continuous stimulation of the groEL ATPase even in the presence of groES (Fig. 5a).

Together, these results suggest that groES couples the groEL ATPase with the folding reaction at the surface of the chaperonin. We propose that groES coordinates Mg-ATP-dependent conformational changes of groEL and thereby allows correct folding to occur by a process of stepwise release. The groES protein prevents the 'wholesale' dissociation of the substrate protein from groEL which is unproductive for folding.

Conclusions

GroEL stabilizes unfolded proteins in a conformation lacking ordered tertiary structure and, regulated by groES, mediates their ATP-dependent progression to a more compact state. Under physiological conditions, soluble globular proteins may generally be sequestered during folding at the surface of a chaperonin-type machinery.

The conformation of groEL-bound polypeptides resembles that of the 'molten globule' or 'compact intermediate', an early folding state that rapidly interconverts with the fully unfolded form^{29–32}. It is characterized by its containing all or part of the secondary structure of the protein in conjunction with a relatively compact, but unordered, flexible tertiary structure internalizing a 'molten' hydrophobic core. The groEL-stabilized intermediates, DHFR-11 and RHO-11, fulfill these criteria: (1) compared with the fully unfolded state, their tryptophan residues are in a less polar environment and are only partially exposed; (2) their hydrophobic interior is much less densely packed than that of the native conformation as demonstrated by its solvent accessibility and the typical incorporation of the hydrophobic

fluorescent probe ANS; (3) their structural flexibility is reflected in the high sensitivity towards protease. A collapsed-state intermediate of rhodanese folding, transiently stable in 2 M GdmCl, has been described recently^{44,45} which already contains a large part of the secondary structure of the native enzyme. This form of rhodanese exhibits fluorescence properties very similar to those of groEL-bound RHO-I1, including the partial solvent accessibility of the Trp residues. We find that rhodanese partially folded in 2 M GdmCl will efficiently bind to groEL (unpublished results). Similarly, DHFR acquires considerable α -helical content within 15 ms of refolding, followed by a collapse of hydrophobic residues to the interior of the molecule within about 150 ms which leads to the stabilization of β sheets (ref. 46; C. R. Matthews, personal communication). This is much less than the time required for manual mixing in our experiments. Our data do not exclude the possibility that groEL interacts with the substrate protein (for example, a nascent polypeptide chain), while it is still in a more extended conformation. We propose, however, that a molten globule-like pre-folded state is a physiological, early intermediate of chaperonin-mediated folding.

Substrate proteins acquire a more rigid structure while maintaining the interaction with the groEL scaffold. The decrease in fluorescence intensity during conversion of DHFR-I1 to DHFR-I2 is probably due to the formation of the natural environment of Trp 24. A mutant form of human DHFR in which Trp 24 is changed to Phe does not show the marked increase in Trp fluorescence intensity on unfolding⁴². The hydrophobic interior of DHFR-I2 and RHO-I2 is more densely packed than in the molten globule state, which accounts for the reduced accessibility for ANS. Similar folding intermediates with partially ordered protease-resistant structure have been described³².

The structural features of unfolded proteins recognized by groEL are so far unknown. The molten globule still contains unstable secondary structure and is thought to expose hydrophobic patches, resulting in the tendency to aggregate. It is unclear whether groEL interacts with contiguous sequences, as demonstrated for hsp70 (ref. 47), or perhaps with structural elements produced by the spatial arrangement of an early folding intermediate. In any case, it has to be assumed that these bound elements are buried within the folding structure on Mg-ATP-dependent release from groEL. Rebinding might occur if this

internalization is unsuccessful. This would allow structural organization by the sampling of different, progressively more compact conformations. Proteins containing only one or two folding domains may represent the 'folding units' handled by groEL. For larger proteins, independent folding units could reach the chaperonin either directly during synthesis or by transfer through components such as hsp70. This principle seems to apply to unfolded proteins that are imported into mitochondria. During and after membrane translocation, they interact with the hsp70 in the mitochondrial matrix before reaching the chaperonin hsp60 for folding^{48,49}.

Folding at the interacting surface between groEL and the substrate protein would rely on groEL initially binding to two (or more) segments of the protein substrate and releasing them sequentially on ATP hydrolysis. This latter step requires the action of groES, which couples the hydrolysis of ATP with folding at groEL. The groES protein might coordinate the function of (adjacent) parts of the substrate-binding region, belonging to different groEL subunits, to prevent the release of the complete protein in a single step and to give a (sub)domain sufficient time to organize to a more compact structure before rebinding occurs. Rhodanese is fully dependent on such a mechanism for folding, whereas DHFR, at least *in vitro*, is not. Nevertheless, the observations made with DHFR are consistent with the above considerations: in the absence of groES, folding is retarded by increasing the concentration of groEL as a result of cycles of unproductive binding and release. But this is not observed in the presence of groES, probably because the protein is then forced onto a folding pathway that involves a staged release during a single round of interaction.

The amount of ATP hydrolysed by groEL per rhodanese molecule folded is in the range of 5–10% of the total required for synthesis. This suggests that the structural organization of a polypeptide at the surface of groEL requires many rounds of ATP hydrolysis by the groEL 14-mer. The energy of the ATP hydrolysed is probably required to promote conformational changes of groEL resulting in (sequential) release of the bound substrate. The chaperonins are essential for folding *in vivo*, that is, they 'catalyse' protein folding in biological systems. It remains to be seen whether they do so by lowering the activation energy of the productive folding pathway, or exclusively by blocking pathways leading to misfolding and aggregation. □

Received 26 April; accepted 30 May 1991.

- Fischer, G. & Schmid, F. X. *Biochemistry* **29**, 2206–2212 (1990).
- Freedman, R. B. *Cell* **57**, 1069–1072 (1989).
- Ellis, R. J. & van der Vies, S. M. A. *Rev. Biochem.* **60**, 327–347 (1991).
- Ellis, R. J. *Nature* **328**, 378–379 (1987).
- Rothman, J. E. *Cell* **59**, 591–601 (1989).
- Pelham, H. R. B. *Nature* **332**, 776–777 (1988).
- Hartl, F.-U. *Semin. Immun.* (in the press).
- Hemmingsen, S. M. *et al. Nature* **333**, 330–334 (1988).
- Georgopoulos, C., Hendrix, R., Casjens, S. R. & Kaiser, A. D. *J. molec. Biol.* **76**, 45–60 (1973).
- Sternberg, N. *J. molec. Biol.* **76**, 25–44 (1973).
- Hendrix, R. W. *J. molec. Biol.* **129**, 375–392 (1979).
- Bochkareva, E. S., Lissin, N. M. & Girshovich, A. S. *Nature* **336**, 254–257 (1988).
- Goloubinoff, P., Christeller, J. T., Gatenby, A. A. & Lorimer, G. H. *Nature* **342**, 884–889 (1989).
- Van Dyk, T. K., Gatenby, A. A. & La Rossa, R. A. *Nature* **342**, 451–453.
- Lecker, S. *et al. EMBO J.* **8**, 2703–2709 (1989).
- Laminet, A. A., Ziegelhoffer, T., Georgopoulos, C. & Plückthun, A. *EMBO J.* **9**, 2315–2319 (1990).
- Buchner, J. *et al. Biochemistry* **30**, 1586–1591 (1991).
- Barracough, R. & Ellis, R. J. *Biochim. biophys. Acta* **608**, 19–31 (1980).
- Musgrove, J. E., Johnson, R. A. & Ellis, R. J. *Eur. J. Biochem.* **163**, 529–534 (1987).
- Gatenby, A. A. & Ellis, R. J. *A. Rev. Cell Biol.* **6**, 125–149 (1990).
- McMullin, T. W. & Hallberg, R. L. *Molec. cell. Biol.* **8**, 371–380 (1988).
- Reading, D. S., Hallberg, R. L. & Myers, A. M. *Nature* **337**, 655–659 (1989).
- Cheng, M. Y. *et al. Nature* **337**, 620–625 (1989).
- Ostermann, J., Horwich, A. L., Neupert, W. & Hartl, F.-U. *Nature* **341**, 125–130 (1989).
- Cheng, M. Y., Hartl, F.-U. & Horwich, A. L. *Nature* **348**, 455–458 (1990).
- Chandrasekhar, G. N., Tilly, K., Woolford, C., Hendrix, R. & Georgopoulos, C. *J. biol. Chem.* **261**, 12414–12419 (1986).
- Lubben, T. H., Gatenby, A. A., Donaldson, G. K., Lorimer, G. H. & Viitanen, P. V. *Proc. natn. Acad. Sci. U.S.A.* **87**, 7683–7687 (1990).
- Viitanen, P. V. *et al. Biochemistry* **29**, 5665–5670 (1990).
- Pittsyo, O. B., Pain, R. H., Semisotnov, G. V., Zerovnik, E. & Razgulyaev, O. I. *FEBS Lett.* **262**, 20–24 (1990).

- Ewbank, J. J. & Creighton, T. E. *Nature* **350**, 518–520 (1991).
- Kuwajima, K. *Proteins* **6**, 87–103 (1989).
- Creighton, T. E. *Biochem. J.* **270**, 1–16 (1991).
- Kumar, A. A., Blankenship, D. T., Kaufman, B. T. & Freisheim, J. H. *Biochemistry* **19**, 667–678 (1980).
- Matthews, D. A. *et al. J. biol. Chem.* **260**, 381–391 (1985).
- Touchette, N. A., Perry, K. M. & Matthews, C. R. *Biochemistry* **25**, 5445–5452 (1986).
- Waxman, L. & Goldberg, A. L. *Science* **232**, 500–503 (1986).
- Ostoa-Salama, P., Ramirez, J. & Perez-Montfort, R. *Biochim. biophys. Acta* **1041**, 140–152 (1990).
- Jaenicke, R. *Prog. Biophys. Molec. Biol.* **49**, 117–237 (1987).
- Ploegman, J. H. *et al. Nature* **273**, 124–129 (1978).
- Tandon, S. & Horowitz, P. M. *J. biol. Chem.* **261**, 15615–15681 (1986).
- Hohn, T., Hohn, B., Engel, A., Wortz, M. & Smith, P. R. *J. molec. Biol.* **129**, 359–373 (1979).
- Huang, S. *et al. Biochemistry* **28**, 471–478 (1989).
- Semisotnov, G. V. *et al. FEBS Lett.* **224**, 9–13 (1987).
- Tandon, S. & Horowitz, P. M. *J. biol. Chem.* **264**, 9859–9866 (1989).
- Tandon, S. & Horowitz, P. M. *J. biol. Chem.* **265**, 5967–5970 (1990).
- Garvey, E. P., Swank, J. & Matthews, C. R. *Proteins struct. funct. Genet.* **6**, 259–266 (1989).
- Flynn, G. C., Chapell, T. G. & Rothman, J. E. *Science* **245**, 385–390 (1989).
- Kang, P.-J. *et al. Nature* **348**, 137–143 (1990).
- Scherer, P. E., Krieg, U. C., Hwang, S. T., Vestweber, D. & Schatz, G. *EMBO J.* **9**, 4315–4322 (1990).
- Fayet, O., Louran, J. M. & Georgopoulos, C. *Molec. gen. Genet.* **202**, 35–445 (1986).
- Vachereau, A. *Analyt. Biochem.* **179**, 206–208 (1989).
- Jaenicke, R. & Rudolph, R. in *Protein Structure, a Practical Approach* (ed. Creighton, T. E.) 191–222 (IRL, Oxford, 1989).
- Lill, R., Dowhan, W. & Wickner, W. *Cell* **60**, 271–280 (1990).
- Lehrer, S. S. *Biochemistry* **10**, 3254–3263 (1971).

ACKNOWLEDGEMENTS. We thank C. Georgopoulos for providing the groE overproducing strain, P. Horowitz for suggesting rhodanese as a model protein, W. Neupert, W. Fenton and N. Pfanner for discussion and H. Decker for access to the spectrofluorimeter. This work was supported by the Deutsche Forschungsgemeinschaft. R.B. is a fellow of the Alexander von Humboldt-Stiftung.

The Hubble constant from VLA measurement of the time delay in the double quasar 0957+561

David H. Roberts*, Joseph Lehar†,
 Jacqueline N. Hewitt† & Bernard F. Burke†

* Department of Physics, Brandeis University, Waltham,
 Massachusetts 02254, USA

† Department of Physics, 26-335, Massachusetts Institute of Technology,
 Cambridge, Massachusetts 02139, USA

IN the double quasar 0957+561, the first detected example of a gravitational lens¹, a quasar at redshift $z=1.41$ appears as two images separated by ~ 6 arcsec and accompanied by an extended radio source²⁻⁴. The Hubble constant, H_0 , can be estimated from the angular separation and the time delay between the appearance of the same brightness variations in the two images—a method that avoids all the usual uncertainties that attend the estimation of distance by the use of standard candles. From 11 years of Very Large Array (VLA) observations we have found the time delay to be 1.40 ± 0.10 yr, a significantly larger and more robust value than has been obtained from optical measurements^{5,6}. Using a standard lens model⁷ and taking the observed velocity dispersion of luminous matter in the lensing galaxy⁸ to measure the true gravitational potential, we find $H_0 = 46 \pm 14$ (42 ± 14) $\text{km s}^{-1} \text{Mpc}^{-1}$ for cosmic density parameter $\Omega_0=0$ (1), in approximate agreement with Rhee's recent estimate⁸. But if, as is likely, there is dark matter in the lensing galaxy, H_0 could be as high as 69 ± 21 (63 ± 21) $\text{km s}^{-1} \text{Mpc}^{-1}$. The agreement of several completely independent estimates of H_0 gives some confidence that the scale of the Universe is indeed known to about a factor of two.

Even before the discovery of the double quasar it was realized that gravitational lensing offers a new method of determining cosmological parameters⁹. The delay, τ , between the arrival of two images of a single object depends on the difference in length of the paths traversed and on the difference in the gravitational potentials encountered by the wavefronts. In particular, τ is inversely proportional to H_0 , and for 0957+561 is relatively insensitive to the density parameter Ω_0 , the cosmological con-

stant Λ and the mean density of matter along the line of sight^{7,9-13}.

We have been monitoring 0957+561 with monthly observations at the VLA since June 1979, obtaining a total of 80 'snapshot' observations at 5 GHz (wavelength $\lambda=6$ cm) in the three largest configurations of the VLA. A VLA image, taken in the A-configuration, of the double quasar is shown in Fig. 1. The important features are the A and B quasar images, the radio core G of the primary lensing galaxy G1 (located near the B image), and the radio jet and double radio lobes near the A image. Figure 2a shows the 5-GHz light curves of the A and B images. Each flux curve shows a slow linear decline, flattens off, and then brightens. From the first two features, we obtained the preliminary estimates¹⁴ $\tau = (1.35 \pm 0.25)$ yr and $R = 0.69 \pm 0.01$, where $R = m_B/m_A$ is the ratio of the two image magnifications. Now that we have seen upturns in both A and B, we can refine our measurement of the time delay and use it to estimate the Hubble constant.

In the gravitational lens interpretation of 0957+561 the two light curves differ only by an epoch offset τ (where $\tau > 0$ means A leads B) and the magnification ratio R , so a natural way to estimate τ is by correlation analysis. The data are not uniformly spaced, however, and there are occasional gaps of approximately four months when the VLA was in an unsuitable configuration (its most compact). Application of standard cross-correlation techniques to such a data set generates artefacts related to the details of the sampling (J.L., J.N.H., D.H.R. and B.F.B., manuscript in preparation). To overcome such problems, Edelson and Krolik¹⁵ proposed the discrete correlation function (DCF), which provides a correlation estimate without requiring interpolation of the data. After trying several techniques, we chose to derive the time delay from a locally-normalized discrete correlation function (LND CF), a modification of the DCF that allows for the fact that the two light curves are not stationary. The ratio R was determined by a polynomial fitting in which the value of τ derived from the LND CF was used to combine the A and B light curves, and a seventh-order polynomial was fitted to minimize χ^2 . Monte Carlo simulations were used to derive confidence limits for τ and R , taking into account both the noise arising from measurement errors and the systematic effects due to the incomplete sampling (J.L., J.N.H., D.H.R. and B.F.B., manuscript in preparation). From this analysis we obtain

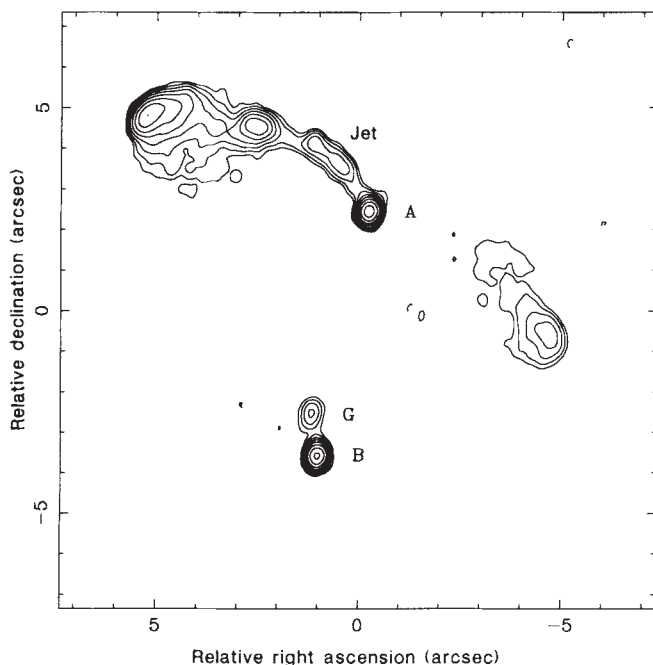


FIG. 1 Radio map of the double quasar 0957+561 made at 5 GHz ($\lambda=6$ cm) using the VLA in the A-configuration, epoch 1982 May 8. The A and B images of the quasar, the radio core G of the primary lensing galaxy, and the radio jet are labelled. The contour levels are $-0.5, 0.5, 1.0, 2.0, \dots, 32$ and 64% of the peak of 31.2 mJy per beam.

$\tau = (1.40 \pm 0.10)$ yr, $R = 0.697 \pm 0.003$. These results are illustrated in Fig. 2*b*, where the B light curve is plotted with the shifted and scaled A light curve, and in Fig. 3, which shows the LNDCF as a function of τ . The short-timescale variability of the quasar images about a single smooth curve could be due to intrinsic variations in the quasar or interstellar scintillations¹⁶. Analysis shows that $\sim 3\%$ flux variations on timescales of several hours should be expected at 5 GHz because of (refractive) interstellar scintillation (J.L., J.N.H., D.H.R. and B.F.B., manuscript in preparation).

Two estimates of the time delay in 0957 + 561 based on optical monitoring have recently been published. Vanderriest *et al.*⁵ find $\tau = (1.136 \pm 0.055)$ yr, and Schild⁶ obtains $\tau = (1.11 \pm 0.03)$ yr (99.9% confidence). Each of these estimates differs from ours by more than two combined standard deviations. Re-analysis of the optical data using the methods described above indicates that the optical uncertainties are larger than those quoted. A new analysis of the optical data (W. H. Press, G. B. Rybicki and J.N.H., manuscript in preparation) yields a value for τ that is compatible with the radio estimate.

Our estimate of the magnification ratio, which applies to the entire light curve, does not agree with the optical estimate from the ratio of emission lines obtained at a single epoch² (0.77 ± 0.03), the optical estimate measurement based on the light curves⁵ (0.97) or the value derived from the very long baseline interferometry (VLBI) structures of the A and B images¹⁷ ($0.64 \pm$

0.02). Temporal variations in the emission lines could account for the first disagreement, and microlensing of the optical B image is a likely explanation for the second^{18,19}. The VLBI estimate is based on the images of the milliarcsecond-scale jets, whose locations are slightly different from those of the centroids of the VLA images, and which would thus have different magnifications. This disagreement could be significant if it indicates a systematic error in the lens model (S. R. Conner, J.L. and B.F.B., manuscript in preparation).

The gravitational lens that forms the multiple images of 0957 + 561 is complex. In addition to the primary lens galaxy G1, it includes the galaxies in the rich cluster of which G1 is the dominant member, as well as any dark matter in that cluster. But numerous observed quantities, including the locations, brightness ratio and relative parities of the A and B images^{3,4,7}, the lack of a detectable third quasar image^{4,20}, and the absence of substantial multiple imaging of most of the extended radio lobes⁴ strongly constrain models of the lens. A detailed discussion of the determination of H_0 from the time delay in 0957 + 561 is given by Falco, Gorenstein, and Shapiro⁷, who show that H_0 , the velocity dispersion σ_v of G1, and τ are related by $H_0 \propto \sigma_v^2 \tau^{-1}$. Sources of uncertainty in the proportionality constant include measurement errors in the VLBI data, the unknown values of Ω_0 and of the clumpiness of the universe, the non-uniqueness of the cluster model, and errors in the detailed model of G1. We have re-evaluated the effect of these

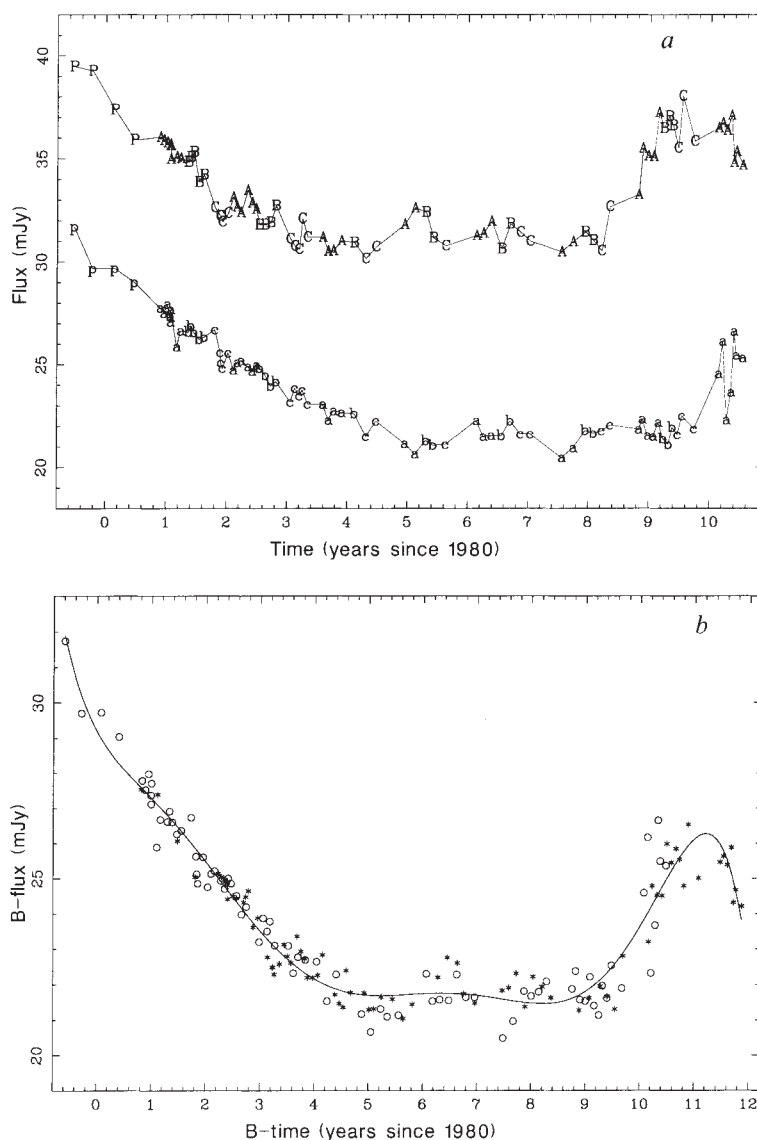


FIG. 2 *a*, VLA light curves of the A and B images of the double quasar 0957 + 561 at 5 GHz. The upper case (lower case) letters denote the A (B) images, and the symbols indicate the VLA configuration in which each observation was obtained (A, B and C are standard VLA configurations, and P denotes a partial [construction] configuration). The zero point of time is 1980. *b*, Light curves of the A and B images, with the A data (asterisks) delayed by 1.40 years and scaled by a factor of 0.697 to match the B data (circles). The smooth curve is the seventh-order Legendre polynomial that best fits this composite light curve at this delay.

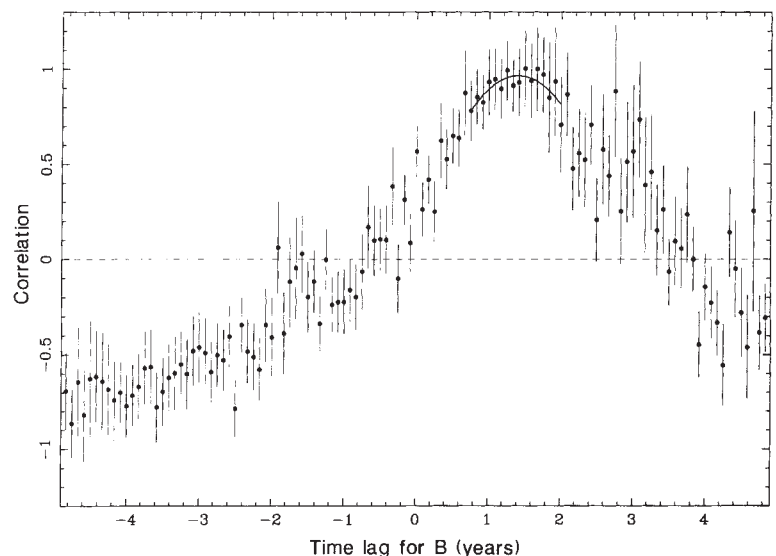


FIG. 3 Locally normalized discrete correlation function (LNDCF) of the A and B light curves of the double quasar 0957+561 as a function of time lag. The smooth curve is a parabola fitted to the peak of the LNDCF.

uncertainties, assuming that the ranges for the proportionality constant due to each error component can be related to standard deviations which propagate in the usual way. Separating out the contribution due to the uncertainty in Ω_0 , we obtain

$$H_0 = \left\{ \begin{array}{c} 97 \pm 20 \\ 90 \pm 21 \end{array} \right\} \left(\frac{\sigma_v}{390 \text{ km s}^{-1}} \right)^2 \times \left(\frac{1 \text{ yr}}{\tau} \right) \text{ km s}^{-1} \text{ Mpc}^{-1} \left\{ \begin{array}{c} \Omega_0 = 0 \\ \Omega_0 = 1 \end{array} \right\} \quad (1)$$

The fractional uncertainty displayed in equation (1) is more conservative by a factor of ~ 2 than that given in reference (7).

The velocity dispersion of G1 has recently been measured by Rhee⁸, who finds $\sigma_v = (303 \pm 50) \text{ km s}^{-1}$. This is in agreement with an estimate of $\sigma_v = (330 \pm 50) \text{ km s}^{-1}$ obtained from the V magnitude of G1 using the Faber-Jackson relation^{7,21}. As their quoted uncertainties are similar, we take the mean value of σ_v , leading (for $\Lambda = 0$) to

$$H_0 = \left\{ \begin{array}{c} 46 \pm 14 \\ 42 \pm 14 \end{array} \right\} \text{ km s}^{-1} \text{ Mpc}^{-1} \left\{ \begin{array}{c} \Omega_0 = 0 \\ \Omega_0 = 1 \end{array} \right\} \quad (2)$$

Detailed study of nearby elliptical galaxies²² has shown that single King models fitted to the central regions cannot account for the observed increase of velocity dispersion with radius, r . A variety of measurements indicate that such galaxies are embedded in r^{-2} potentials; assuming a Hubble brightness profile²³, the optical velocity dispersion underestimates the total velocity dispersion by 50%. If such a model applies to G1, then $H_0 = (69 \pm 21) [(63 \pm 21)] \text{ km s}^{-1} \text{ Mpc}^{-1}$ for $\Omega_0 = 0$ [$\Omega_0 = 1$], permitting H_0 to be as large as $90 \text{ km s}^{-1} \text{ Mpc}^{-1}$.

Estimating H_0 with the double quasar probes the universe to much greater redshifts than do traditional methods, so it is necessary to examine any systematic uncertainties due to the effect of the unknown value of Λ . In 0957+561 the combination of distances used to calculate the properties of the lens varies by no more than 5% for $\Lambda/\Lambda_{\text{crit}} \leq 0.9$ ($\Omega = 0.1$)¹³, and depends on the nature of the model components involved (for example, point masses or King models). We estimate that this results in a similar fractional uncertainty in the resulting value of H_0 , but as little is known about the true value of Λ , we have not included this uncertainty in our error estimates.

The value of the Hubble constant has been the subject of an extended debate between two schools, one preferring H_0 in the range $40\text{--}60 \text{ km s}^{-1} \text{ Mpc}^{-1}$ (a larger, older universe), and a second insisting on $80\text{--}100 \text{ km s}^{-1} \text{ Mpc}^{-1}$ (smaller and younger)²⁴. At present our estimate is not accurate enough to distinguish between these two possibilities. Better observations of the velocity dispersion and morphology of G1 and more

detailed VLA images of the B-G region offer the greatest opportunities to improve the estimate of H_0 from 0957+561.

Several independent techniques have now been used to estimate the Hubble constant. Methods based on standard candles measure luminosity distance, comparison of the transverse and radial expansion rates of supernova shells measures proper motion distance²⁵, and the time delay in a gravitational lens measures angular size distance. All now yield estimates of H_0 in the range $50\text{--}100 \text{ km s}^{-1} \text{ Mpc}^{-1}$, suggesting that the scale of the universe is indeed known to within a factor of two. Finally, we note that detection of identical but time-delayed light curves for the A and B images is additional proof of the gravitational lens interpretation of the double quasar, showing that the object quasar at $z = 1.41$ lies beyond the lensing galaxy at $z = 0.36$.

Note added in proof. Kochanek²⁶ has suggested an alternative model for 0957+561 which gives values for H_0 that are approximately 50% larger than those derived from the model of Falco, Gorenstein and Shapiro. $\square$

Received 9 April; accepted 23 May 1991.

- Walsh, D., Carswell, R. F. & Weymann, R. J. *Nature* **279**, 381–384 (1979).
- Young, P. J. *et al. Astrophys. J.* **241**, 507–520 (1980).
- Young, P. J. *et al. Astrophys. J.* **244**, 736–755 (1981).
- Roberts, D. H. *et al. Astrophys. J.* **293**, 356–369 (1985).
- Vanderriest, C. *et al. Astr. Astrophys.* **215**, 1–13 (1989).
- Schild, R. E. *Astrophys. J.* **100**, 1771–1776 (1990).
- Falco, E. E., Gorenstein, M. V. & Shapiro, I. I. *Astrophys. J.* **372**, 364–379, *Center for Astrophysics Preprint No. 3103* (1991).
- Rhee, G. *Nature* **350**, 211–212 (1991).
- Riess, S. *Mon. Not. R. astr. Soc.* **132**, 101–111 (1966).
- Dyer, C. C. & Roeder, R. C. *Astrophys. J.* **241**, L133–L136 (1980).
- Borgeest, U. & Riess, S. *Astr. Astrophys.* **141**, 318–322 (1984).
- Alcock, C. & Anderson, N. *Astrophys. J.* **302**, 43–51 (1986).
- Fukugita, M., Futamase, T. & Kasai, M. *Mon. Not. R. astr. Soc.* **246**, 24P–27P (1990).
- Lehar, J., Hewitt, J. N. & Roberts, D. H. *Gravitational Lenses* (eds Moran, J. M., Hewitt, J. N. & Lo, K. Y.) 84–89 (Reidel, Dordrecht, 1989).
- Edelson, R. A. & Krolik, J. H. *Astrophys. J.* **333**, 646–659 (1988).
- Rickett, B. J. *Ann. Rev. Astr. Astrophys.* **28**, 561–605 (1990).
- Gorenstein, M. V. *et al. Astrophys. J.* **334**, 42–58 (1988).
- Young, P. J. *Astrophys. J.* **244**, 756–767 (1981).
- Schild, R. E. & Smith, R. C. *Astr. J.* **101**, 813–817 (1991).
- Bonometti, R. J. thesis, Massachusetts Institute of Technology (1985).
- Faber, S. M. & Jackson, R. E. *Astrophys. J.* **204**, 668–683 (1976).
- Kent, S. M. *Evolution of the Universe of Galaxies* (ed. Kron, R. G.) 109–118 (Astronomical Society of the Pacific, San Francisco, 1990).
- Gott, J. R. III *Ann. Rev. Astr. Astrophys.* **15**, 235–266 (1977).
- van den Bergh, S. *The Extragalactic Distance Scale* (eds van den Bergh, S. & Pritchet, C. J.) 375–377 (Astronomical Society of the Pacific, San Francisco, 1988).
- Bartel, N. IAU Symp. No. 129, *The Impact of VLBI on Astrophysics and Geophysics* (eds Reid, M. J. & Moran, J. M.) 175–184 (Reidel, Dordrecht, 1988).
- Kochanek, C. S. *Astrophys. J.* (in the press).

ACKNOWLEDGEMENTS. We thank B. Clark, M. Perley, D. Wunker and P. Greenfield for help during this project. E. Falco, M. Fukugita, C. Lonsdale, R. Schild, E. Turner and J. Wardle for discussions, G. Rhee for communicating his result before its publication and J. Moran for making Center for Astrophysics Facilities available for the calibration of some of the data. This work was supported by NSF at Brandeis (D.H.R.), MIT (B.F.B.) and Haystack Observatory. J.N.H. is an Alfred P. Sloan Fellow and a David and Lucile Packard Fellow. The NRAO is operated by Associated Universities, Inc., under cooperative agreement with the NSF.

Coherent backscatter model for the unusual radar reflectivity of icy satellites

Bruce Hapke* & David Blewett†

Department of Geology and Planetary Science, University of Pittsburgh, Pittsburgh, Pennsylvania 15260, USA

RADAR is a powerful technique for probing the surfaces and subsurfaces of Solar System bodies. Inner Solar System bodies reflect radar in an almost specular fashion, with low reflectivity and little polarization. The icy satellites of Jupiter, by contrast, show high reflectivities, diffuse scattering laws and unusual polarization properties¹: compared with what would be expected for specular reflection, there is 1.5 times as much power reflected in the unexpected sense of circularly polarized radar as in the expected sense, and half as much power in the unexpected as in the expected sense of linear polarization. According to the coherent backscatter model², most of the received power from icy satellites is multiply reflected by particles about a wavelength in size located randomly under the surface of the regolith. We have constructed a laboratory analogue of this model by reflecting laser light off polystyrene beads suspended in water, and find that this model reproduces the unusual polarization ratios observed in the radar data. This implies that the regoliths of icy satellites are weakly absorbing matrices of small refractive index containing imbedded scatterers separated by distances of the order of a wavelength. No structures of special shape or other geometrical or optical properties are required.

The radar scattering properties of the icy satellites have been a puzzle for over a decade, and as it will be at least another decade before any spacecraft will land on any icy satellite, remote sensing remains the only way to obtain information about their surfaces. Of the unusual radar reflection properties of the icy satellites, the large circular polarization ratios have been the most baffling. The success of simple specular reflection models for the inner Solar System bodies has led to attempts to explain the signals from the icy bodies by postulating reflections from unusual surface structures (see ref. 1), but these have been considered geologically unlikely.

Several workers²⁻⁵ have pointed out that the large, diffuse reflectances from the icy bodies can be understood if most of the signal occurs by multiple scattering from sub-surface discontinuities in refractive index, that is, from scatterers imbedded in a weakly absorbing medium. One of us² has recently published such a model, which accounts quantitatively for both the amplitude and shape of the radar signals. Intrinsic to this particular model is a phenomenon known as coherent backscatter, which not only increases the reflectance, but also has unusual polarization properties. This phenomenon has been studied in a limited way both experimentally and theoretically. But because of the formidable mathematics involved, a complete, quantitative, vector theory of the effect has not yet been published (see refs 2 and 6 for more thorough discussions and additional references).

In a medium containing scatterers separated by distances of the order of the wavelength, portions of waves that are multiply scattered between particles and that happen to traverse the same path, but in opposite directions, combine coherently in the backscatter direction. This produces a narrow peak in the angular scattering function centred at zero phase angle, with an intensity that may be nearly twice that of the continuum. Outside the backscatter peak, multiple scattering tends to randomize both the directions and polarizations of the photons. Within the

peak, however, coherent effects tend to preserve the initial sense of polarization. Hence, if the incident radiation is linearly polarized, a large fraction of the backscattered radiation is polarized parallel to the incident direction of polarization. If the illumination is circularly polarized, a large fraction of the light in the backscatter peak is polarized in the same sense as the incident radiation.

Linear polarization ratios similar to those in the radar returns from icy satellites have been observed in experimental investigations of coherent backscatter. Circular polarization ratios, however, have not been reported. We have measured the relative intensities, and linear and circular polarization ratios, in the coherent backscatter peak in a laboratory analogue experiment at optical wavelengths. The scattering medium was a 10% suspension of 0.497- μm polystyrene spheres in water (Duke Scientific). This was placed in a cup and illuminated by light of wavelength 0.633 μm from a helium-neon laser. After leaving the laser the light passed through a polarizing filter with its axis oriented parallel to the plane containing the incident beam, the normal to the target surface and the direction to the detector. It was then reflected off a front-surface mirror, from which it illuminated the medium at an angle of 5° from the vertical.

The light scattered from the target was observed with a moveable photomultiplier, in front of which was a rotatable, polarizing filter. A quarter-wave plate could be placed over the target to convert the incident light from linear to circular polarization, and vice versa for the scattered light. The source was kept fixed,

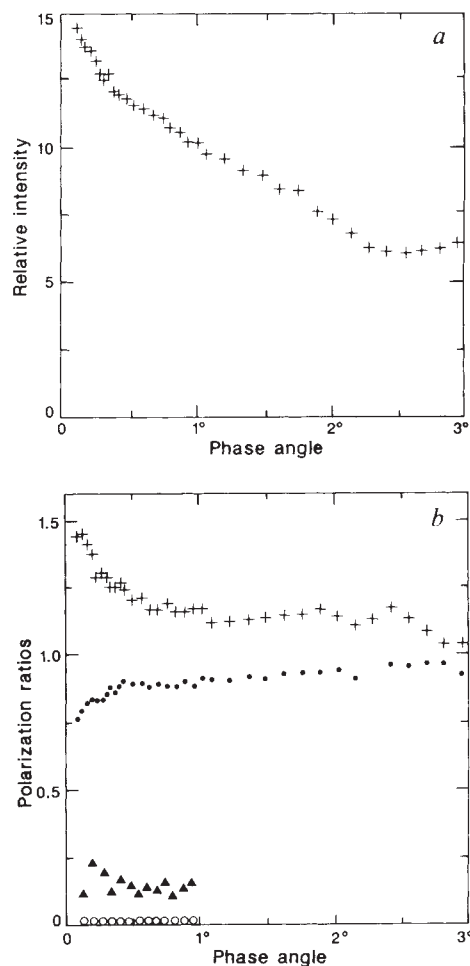


FIG. 1 Relative intensity (a) and circular (+) and linear (O) polarization ratios (b) plotted against phase angle of light scattered from a suspension of polystyrene spheres in water. Also shown in the lower part of b are the linear (O) and circular (▲) polarization ratios of light specularly reflected from the surface of a silicate glass.

* To whom correspondence should be addressed at: 321 Old Engineering Hall, University of Pittsburgh, Pittsburgh, Pennsylvania 15260, USA.

† Present address: Planetary Geosciences Division, University of Hawaii, Honolulu, Hawaii 96822, USA.

and the intensity and linear and circular polarizations of the scattered light were measured as a function of direction to the detector. The aperture sizes were such that the angular resolution and minimum phase angle were both $\sim 0.09^\circ$. Radar observations are made exactly at zero phase angle; unfortunately, our apparatus did not allow measurements at zero phase, but the response there can be inferred by extrapolation from larger angles.

The results of the measurements are shown in Fig. 1. Figure 1a displays the relative intensity against phase angle. The broad rise, $\sim 2.5^\circ$ wide, is the glory, which is a phenomenon associated with scattering by spheres. The coherent backscatter peak is much narrower, with a half-width at half maximum of 0.5° centred on zero phase angle, and is superimposed on the glory. Although the glory has been discussed by many authors^{7,8} it is only partially understood. Here we will not discuss the glory further because it is unrelated to the coherent backscatter.

Figure 1b shows the polarization ratios. The circular polarization ratio is the ratio of intensity of light scattered in the same sense as the incident light to that in the opposite sense. At large phase angles the ratio is close to one, as expected for light that has been randomized by multiple scattering. In the backscatter peak, however, the ratio rises abruptly, and extrapolation from the plot indicates that it is ~ 1.5 at zero phase angle.

The linear polarization ratio, also shown in Fig. 1b, is the ratio of intensity scattered in the opposite sense as the incident light to that in the same sense. This ratio is also close to the expected value of one at large phase angles, but as the backscatter direction is approached, the ratio drops towards a level around 0.7. This is consistent with other published measurements of 0.46–0.70.

The points at the bottom of Fig. 1b demonstrate that these ratios are not caused by some anomalous instrumental effect. This part of the figure shows the linear and circular polarization ratios measured from the smooth, slightly rippled surface of a piece of dark welding glass. The signal consists almost entirely of light specularly reflected from the surface of the glass. As expected, there is no sign of a coherent backscatter peak, and both polarization ratios remain small at all angles.

Our measurements of the polarization ratios in the coherent backscatter peak, taken together with the scalar diffuse scattering model published previously², show that the coherent backscatter model can reproduce quantitatively all of the hitherto puzzling features of the radar signals from icy satellites. Although our experiment was done at optical frequencies, there is no reason why it should not apply at radio frequencies. It must also be emphasized that coherent backscatter occurs in media made up of irregularly shaped particles as well as in those containing spherical particles. All that this model requires is that a regolith consist of a matrix of small complex refractive index, in which are embedded scatterers separated by distances of the order of a wavelength. As water ice occurs on the surfaces of these bodies⁹, a good candidate for the matrix material is water frost. The embedded scatterers could be voids, chunks of solid ice or silicate rocks.

The angular width of the coherent backscatter peak is roughly $\lambda/2\pi L$, where λ is the wavelength and L is the transport mean free path of the photons⁶. Assuming that the matrix material is ice, an upper limit to L is the absorption length in frost (a few hundred wavelengths). All radar measurements at present have been monostatic, but bistatic radar observations would allow L to be measured. Further laboratory, theoretical and observational work may allow limits to be placed on the sizes or other properties of the sub-surface scatterers.

The regoliths of these jovian satellites may be contrasted with that of the Moon, which consists of a silicate powder matrix containing embedded rocks and boulders. Apart from composition, one main difference seems to be that the lunar sub-surface scatterers are separated by distances that are not small compared with the photon absorption length in the powder (of the order

of tens of wavelengths¹⁰) so that most of the returned signal is by quasi-specular reflection from the surface, rather than from the scatterers. The lunar regolith is generated by meteorite impacts; the dissimilar structures of the icy regoliths indicates that they may have a different origin.

Because coherent backscatter does not require structures of any special form, it is likely to be important in the radar retro-reflection from many other objects in the solar system that are strong scatterers of radio waves, including Io¹¹, Titan¹², the rings of Saturn¹² and the bright ejecta blankets around fresh craters observed on Venus by the Magellan spacecraft. Also, the source of illumination need not be coherent; consequently, it may play a part in the narrow opposition-effect peaks observed at optical frequencies on some of the icy satellites^{14,15}. □

Received 20 February; accepted 20 May 1991.

1. Ostro, S. in *Satellites of Jupiter* (ed. Morrison, D.) 213–236 (University of Arizona Press, Tucson, 1982).
2. Hapke, B. *Icarus* **88**, 407–417 (1990).
3. Goldstein, R. & Green, R. *Science* **207**, 179–180 (1990).
4. Hagfors, T., Gold, T. & Jerkic, H. *Nature* **315**, 637–640 (1990).
5. Ostro, S. & Shoemaker, E. *Icarus* **85**, 335–345 (1990).
6. MacKintosh, F. & John, S. *Phys. Rev.* **B37**, 1884–1897 (1988).
7. Kerker, M. *The Scattering of Light and Other Electromagnetic Radiation* (Academic, New York, 1969).
8. Bohren, C. & Huffman, D. *Absorption and Scattering of Light by Small Particles* (Wiley, New York, 1983).
9. Clark, R. & McCord, T. *Icarus* **41**, 323–339 (1980).
10. Gold, T., O'Leary, B. & Campbell, M. *Proc. 2nd Lunar Planet. Sci. Conf.*, 2173–2181 (1971).
11. Ostro, S., Resema, K., Campbell, D., Chandler, J. & Shapiro, I. *Bull. Am. Astr. Soc.* **22**, 1109 (1990).
12. Muhleman, D., Grossman, A., Butler, B. & Siade, M. *Science* **248**, 975–980 (1990).
13. Cuzzi, J. & Pollack, J. *Icarus* **33**, 233–262 (1978).
14. Brown, R. & Cruikshank, D. *Icarus* **55**, 83–92 (1983).
15. Domingue, D., Hapke, B., Lockwood, G. & Thompson, D. *Icarus* **90**, 30–42 (1991).

ACKNOWLEDGEMENTS. This work is supported by the Planetary Geology and Geophysics Program, Office of Space Science and Applications of NASA.

Artificial 'olfactory' images from a chemical sensor using a light-pulse technique

I. Lundström, R. Erlandsson, U. Frykman, E. Hedborg, A. Spetz, H. Sundgren, S. Welin & F. Winquist

Laboratory of Applied Physics, University of Linköping, S-58183 Linköping, Sweden

THERE is much interest in the use of chemical sensor arrays, in conjunction with pattern-recognition routines, for developing artificial olfactory devices—'electronic noses'—which can characterize the chemical composition of gas mixtures^{1–5}. Here we describe a technique that uses a continuous sensing surface and a detection method involving a scanning pulsed light source, to generate images that represent a fingerprint of the gases detected. The detector is a large-area field-effect device with a number of different catalytic metals constituting the detecting surface (the device's active gate)^{6,7}. A pulsed light beam scanned across this surface generates a photocapacitive current that varies with the value of the surface potential^{8,9}. A continuous sensing surface of this type provides information that would require an array of hundreds of discrete sensors. The technique also provides a new means of studying the coupling between the electronic properties of catalytic metals and chemical reactions taking place on their surfaces.

Chemical sensors based on field-effect structures with gates of catalytic metals are sensitive to molecules such as hydrogen, ammonia, hydrogen sulphide, ethylene and ethanol^{6,7,10}. The selectivity patterns of the sensors depend on the operating temperature, and on the type and thickness of the catalytic metal used. These sensors have been used in discrete sensor arrays to predict the concentration of H₂ and NH₃ in (four-component)

gas mixtures in air^{11,12}. In sensors with thin discontinuous metal gates, changes in the surface potential of the semiconductor are induced by potential changes on the metal surface, at the metal-insulator interface and on the exposed insulator^{6,7}. The surface potential of the semiconductor can be measured with a semi-transparent metal contact and a chopped light beam, as the light pulses produce a photocapacitive current in the metal-insulator-semiconductor structure. If the light is scanned along the surface, a map of the surface potential can be obtained. This technique has been used to detect defects in semiconductor surfaces^{8,9}. A chopped-light-beam technique has also been used to detect pH changes outside an insulator-semiconductor structure through the influence of the voltage drop across the insulator-electrolyte interface on the surface potential of the semiconductor¹³; this has been used to monitor the metabolism of living cells¹⁴. Here we point out that a modified scanning-light-pulse technique is well suited to detect changes in the electronic properties of thin (discontinuous) catalytic metal films caused by chemical reactions on the film. Chemical reactions can thus be studied on a catalytic metal surface with a resolution given by the size of the light spot. This could be used to study the development of chemical oscillations on the catalytic metal surface at atmospheric pressures. Here, however, we wish to stress another possibility, namely the direct formation of patterns or images of gas mixtures by the use of a large-area surface having differences in the catalytic selectivity and activity along the surface.

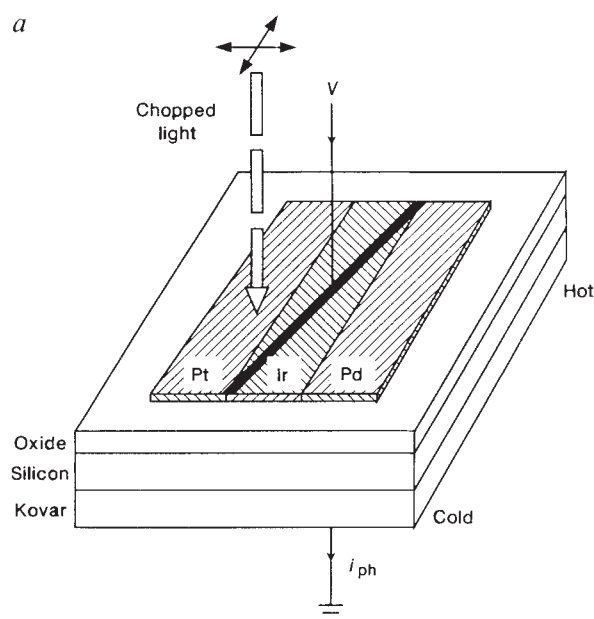
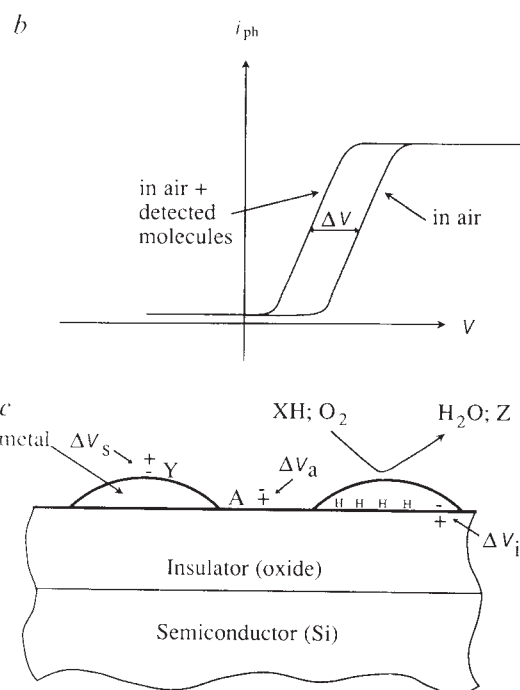


FIG. 1 *a*, Schematic of the experimental arrangement. A metal/silicon dioxide/silicon structure with a total metal area of $4 \times 6 \text{ mm}^2$ was fabricated by standard methods. The oxide was thermally grown on p-type Si ($7-13 \Omega \text{ cm}^{-1}$) in dry oxygen at $1,200^\circ \text{C}$ to a thickness of 100 nm. The metal gate consisted of three different thin catalytic metal layers, as labelled in the figure. The metal layers were prepared by thermal evaporation through metal masks. The thickness of the metal layers was $\sim 6 \text{ nm}$. A thick contact strip (Pd) was evaporated across the structure (black line) and the structure was pressed onto a Kovar plate. One end of the sample was in contact with a heating element (a resistor) and the other was hanging in the air. The temperature of the hot and cool ends were roughly 180°C and 110°C respectively (determined by calibrating the electrical properties of the structure as a function of temperature). A photocapacitive current was produced with a chopped light beam, scanned mechanically over the sample (see Fig. 2). The diameter of the light spot was $\sim 0.3 \text{ mm}$. The current depends on the width of the depletion layer in the semiconductor (or the surface potential

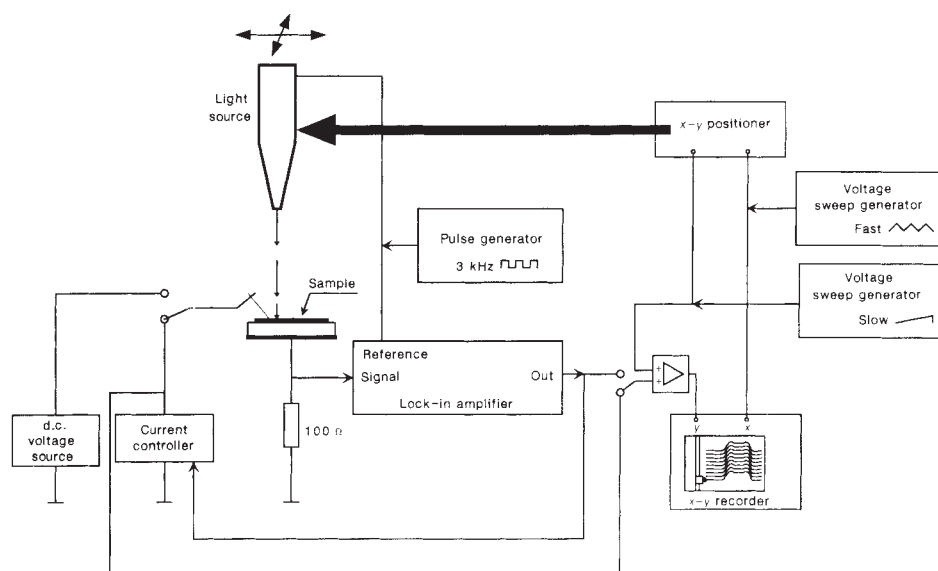
Our samples consist of metal/silicon dioxide/silicon structures with large-area metal contacts. Figure 1*a* shows the structure schematically. By scanning the chopped light beam over the metal surface with and without the molecules to be detected present, we can measure a voltage difference $\Delta V(x, y)$, defined in Fig. 1*b*. If the properties of the catalytic metal vary along the surface, for example because of differences in composition and operating temperature, then a map is obtained which in principle will be different for each gas composition. This we term an artificial 'olfactory' image.

Figure 2 shows a schematic of the experimental setup. One special detail is the large scan size, as the metal surface was $4 \times 6 \text{ mm}^2$. We therefore attached the light source, a frequency-modulated light-emitting diode and a focusing lens to the pen holder of an x-y recorder, which was used as the scanner. To overcome the fact that the work function of the metals used are different and that the location of the photocurrent-voltage $i_{ph}(V)$ curve (Fig. 1*b*) also depends on temperature, a feedback circuit was used to monitor the d.c. voltage necessary to give a constant preset photocapacitive current. In this way the chemically induced shift of the $i_{ph}(V)$ curve, ΔV , could be monitored directly. The connection between the potential changes and ΔV (Fig. 1*c*) has not been thoroughly worked out. It has been suggested⁷ that the potential change on the metal surface, ΔV_s , is capacitively coupled to the semiconductor surface through the voids in the thin film. The measured ΔV will in general have



of the semiconductor). The technique therefore measures local variations in surface potential. *b*, Schematic example of photocapacitive current at a given location of the light spot for p-type silicon. The dashed curve illustrates the influence of a work function or rather a potential change between the metal and the semiconductor, ΔV , caused by adsorption and/or catalytic reactions on the surface. *c*, Schematic diagram of a thin catalytic metal film on an insulator surface and its interaction with gas molecules. If the metal is sufficiently thin it will have openings in which the insulator is exposed, and a surface potential change of the semiconductor can then be caused by hydrogen atoms at the metal-insulator interface (ΔV_i), by adsorbates (A) on the insulator (ΔV_a) or by reaction intermediates or adsorbates (Y) on the metal surface (ΔV_s). ΔV_s is effectively capacitively coupled to the semiconductor surface through the openings of the metal film⁷. $\text{XH} + \text{O}_2 \rightarrow \text{H}_2\text{O} + \text{Z}$ denotes a dehydrogenating reaction of XH in air producing water molecules and other products, Z.

FIG. 2 Schematic of the scanning light-pulse arrangement. The power to the LED was modulated at a frequency of 3 kHz. The light source was moved by an x-y recorder driven by signal generators as indicated. Each x-scan (along the metal strips) took 10 s; a full map (a full y-scan) took 200 s. Two modes of operation were possible: direct measurement of the photocapacitive current, i_{ph} , at a given d.c. bias, or the monitoring of the d.c. bias necessary to give a preset i_{ph} . The diagram shows the switches set in the second (current-controlled) mode. The sample was subjected to a flow of synthetic air (20% O₂ in argon), supplied by an open tube pointing at the cold end of the sample (not shown), with the test gases mixed in as needed by a computerized gas-mixing system.



contributions from ΔV_i , ΔV_s and ΔV_a (see legend to Fig. 1c).

Figure 3a-c shows examples of three-dimensional maps of $\Delta V(x, y)$ for three different molecules in air. We obtained the maps with the arrangement in Fig. 1a. The pictures were produced with commercial image-processing software (Super 3D, Silicon Beach Software Inc.), in which the amount of shadowing was determined by illuminating the maps with simulated light sources. Although these maps contain all information obtained from the measurement, it is instructive to produce pictures from them by assigning different colours or grey tones to different

ΔV . For this we used a computer program originally developed for scanning force microscopy¹⁵. Figure 3d-f shows grey-scale-coded 'top views' of the data in Fig. 3a-c. These pictures clearly show the different selectivities of the three metals used and the influence of temperature on their sensitivity to the three different molecules. The chemistry on the surface has thus been translated into a visual impression.

Our results indicate that with sufficiently large sensor matrices or continuous sensor surfaces, as in our example, conventional image processing can be used to identify odours. We believe

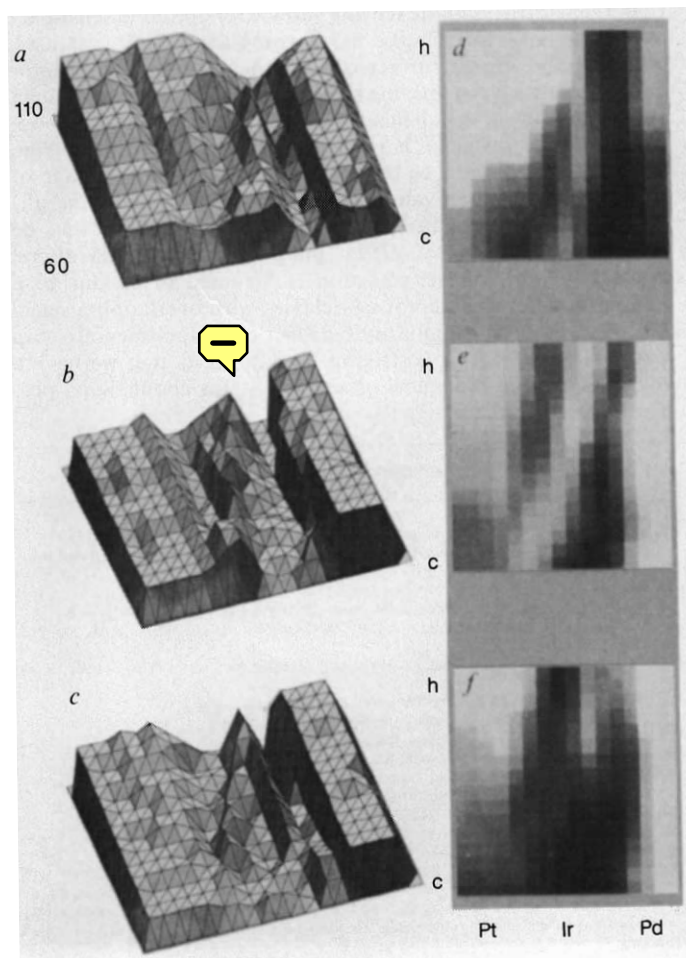


FIG. 3 Shift of the photocapacitive current-voltage curve induced by 100 p.p.m. of *a*, ammonia, *b*, hydrogen and *c*, ethanol in air, as a function of position on the test structure in Fig. 1a. *h* stands for the hot end, *c* for the cold end of the test structure. $\Delta V(x, y)$ was obtained by recording the voltage necessary to yield a given photocapacitive current before and after exposure to the molecules. This is a simple way to measure the parallel shift of the photocurrent-voltage characteristics caused by the molecules. The observed values of ΔV (in mV) are indicated in *a*. In these preliminary studies the shifts were calculated for a matrix consisting of 18×18 points, which was the basis for the plots in *a-c*. The shadowing in the drawings is related to the spatial derivative of $\Delta V(x, y)$. $\Delta V(x, y)$ was also grey-scale coded¹⁵ to provide a 'top view' of the maps in (*a-c*). Black corresponds to $\Delta V < 25$ mV and white to $\Delta V > 65$ mV. The other shades were uniformly distributed within this interval. In the present experiments ΔV was determined with a resolution of 5 mV. In Fig. 2d-f, we filtered $\Delta V(x, y)$ at each point by using the average ΔV for the point (*x, y*) and its eight nearest neighbours. After the filtering the points closest to the edges were disregarded; the pictures in *d-f* therefore contain 16×16 data points. The pictures represent what can be called 'artificial olfactory images', a different picture for each gas combination. We have chosen three very simple mixtures for illustrative purposes: in *d-f* we observe (1) the high and temperature-independent sensitivity of palladium to hydrogen; (2) the high sensitivity to ammonia of hot platinum and iridium; (3) the temperature-dependent sensitivity of platinum to ethanol; (4) the low sensitivity to ethanol and hydrogen of iridium. The thick contact strip is visible as a diagonal line in the hydrogen picture (*e*).

that presenting chemical information in the form of pictures will be of interest when combined with computer graphics and pattern recognition. Established programs and procedures for the interpretation of visual images could then be used for the further processing of the data. This opens up new possibilities for chemical sensor technologies (electronic 'noses') where a (two-dimensional) gradient in selectivity can be produced along a surface. One possibility would be to use arrays of metal-gate field-effect devices, with one data point (one 'pixel') obtained from each field-effect device. The continuous physical method presented here is, however, very powerful for basic studies and for formation of images with high resolution. If each data point were generated by a sensor, our results would require an array of 324 sensors. Our sensor surface therefore represents a large sensor array with a certain redundancy (partly because of the averaging explained in Fig. 3, partly because there are several 'pixels' across each metal strip) to provide better statistics in the pattern formation. As it is also possible to produce gradients in the properties of insulator surfaces, we suggest that images could also be created from the ionic composition of an electrolyte by using an array of ion-sensitive field-effect devices or a suitably treated continuous insulator surface.

The method described could also be used to investigate (large-area) catalytic metal surfaces under inhomogeneous reaction conditions. This is not only of interest for sensor development where different metal layer thicknesses, composition and operation temperatures can be investigated in the same gas mixture; it should also be possible to study spillover effects and spatial features of chemical oscillations on catalytic surfaces¹⁶ with a focused and pulsed light beam.

Our method could be developed much further. For example, we have not yet optimized the size of the light beam. We have chosen to expose the sensing surface to very simple gas mixtures to demonstrate the principles of the proposed technique. One of the goals is to tailor the sensing surface for optimum sensitivity and selectivity to identify more complicated gas mixtures. Although the sensing surface is very different from the receptor cells in olfactory systems, the signal processing possible through the formation of visual images resembles current ideas about olfaction. For instance, it is suggested¹⁷ that the signals from the receptor cells can be transformed into a "spatial image of the odor with analog values at discrete points (the glomeruli) on a surface. The further processing of this image can thus be understood in terms of digital image processing and digital filtering". Furthermore, olfaction is assumed to be due to a limited number of receptor-cell classes with overlapping selectivity patterns. The similarity between our experimental setup and olfactory systems is striking. This suggests that we have a lot to learn from biological olfactory systems about signal processing of the sensor signals. □

Received 23 January; accepted 29 April 1991.

- Carey, W. P., Beebe, K. R., Kowalski, B. R., Illman, D. L. & Hirschfeld, T. *Analyt. Chem.* **58**, 149–153 (1986).
- Müller, R. & Lange, E. *Sensors Actuators* **9**, 39–48 (1986).
- Kaneyasu, M., Ikegami, A., Arima, H. & Iwanaga, S. *IEEE Trans. Components, Hybrids Manufact. Technol. CHMT-10*, 267–273 (1987).
- Shurmer, H. V. *IEEE Proc. G* **137**, 197–204 (1990).
- Gardner, J. W., Hines, E. L. & Wilkinson, M. *Meas. Sci. Technol.* **1**, 446–452 (1990).
- Lundström, I., Armgarth, L.-G. & Pettersson, L.-G. *CRC Critical Rev. Solid St. Mater. Sci.* **15**, 201–278 (1989).
- Lundström, I., Spetz, A., Winquist, F., Ackelid, U. & Sundgren, H. *Sensors Actuators B* **1**, 15–20 (1990).
- DiStefano, T. H. & Viggiano, V. M. *IBM J. Res. Dev.* **18**, 94–99 (1974).
- Engström, O. & Carlsson, A. *J. appl. Phys.* **54**, 5245–5251 (1983).
- Winquist, F. & Lundström, I. *Sensors Actuators* **12**, 255–261 (1987).
- Sundgren, H. et al. *Sensors Actuators B* **1**, 473–476 (1990).
- Sundgren, H., Winquist, F., Lukkari, I. & Lundström, I. *Meas. Sci. Technol.* **2**, 464–469 (1991).
- Hafeman, D. G., Parce, J. W. & McConnell, H. M. *Science* **240**, 1182–1185 (1988).
- Parce, J. W. et al. *Science* **246**, 243–247 (1989).
- Erlandsson, R., Wigren, R. & Olsson, L. *Microsc. Microanal. Microstruct.* (in the press).
- Rotermund, H. H., Engel, W., Kordes, M. & Ertl, G. *Nature* **243**, 355–357 (1990).
- Schild, D. *Biophys. J.* **54**, 1001–1011 (1988).

ACKNOWLEDGEMENTS. This work was supported by grants from the Swedish National Board for Technical Development and from the Center for Industrial Information Technology, University of Linköping.

Composite membranes from photochemical synthesis of ultrathin polymer films

Chao Liu and Charles R. Martin*

Department of Chemistry, Colorado State University, Fort Collins, Colorado 80523, USA

THERE has recently been a resurgence of interest in synthetic membranes and membrane-based processes^{1–12}. This is motivated by a wide variety of technological applications, such as chemical separations^{1–7}, bioreactors and sensors^{8,9}, energy conversion^{10,11} and drug-delivery systems¹². Many of these technologies require the ability to prepare extremely thin, defect-free synthetic (generally polymeric) films, which are supported on microporous supports to form composite membranes. Here we describe a method for producing composite membranes of this sort that incorporate high-quality polymer films less than 50-nm thick. The method involves interfacial photopolymerization of a thin polymer film on the surface of the microporous substrate. We have been able to use this technique to synthesize a variety of functionalized ultrathin films based on electroactive, photoactive and ion-exchange polymers. We demonstrate the method here with composite membranes that show exceptional gas-transport properties.

The US Department of Energy has recently published an extensive report on membrane separations^{6,7} which identifies several research priorities in this area. One priority is development of methods for forming membranes with defect-free 'skins' (the films that accomplish the desired separation) that are less than 50 nm thick. Such extraordinarily thin, defect-free films would be useful in other applications including sensors^{8,9} and drug delivery systems¹².

One approach is to form a composite by bonding a thin film of a chemically selective material to the surface of a microporous support membrane^{2–4}. Such composite membranes can provide good mechanical properties, high flux and good chemical selectivity. This combination of properties (essential to many applications of thin-film technology) would be nearly impossible to achieve with a homogeneous membrane^{2–4}. The challenge, then, is to develop procedures for forming composite membranes with defect-free, chemically selective films less than 50 nm thick^{6,7}.

The method described here accomplishes these goals. It entails photochemical synthesis of an ultrathin polymer film on the surface of a microporous support membrane (Fig. 1). The support membrane is placed on a filter paper which is saturated with a solution of the desired monomer(s). This solution is drawn into the pores, rises to the upper surface of the membrane through capillary action and covers the surface with a thin solution film. The surface is then irradiated with a xenon-arc lamp which initiates polymerization of the ultrathin polymer film. The rate of polymerization is enhanced by adding a photoinitiator (benzoin methyl ether)¹³ to the gas phase above the membrane (Fig. 1).

The photoinitiator induces radical polymerization¹³. Thus, this method should be applicable to many monomers. To illustrate this point, we have prepared polymer films from many monomeric starting materials; Table 1 gives examples. Interesting (and useful) functionalized co- and terpolymers can be prepared, including electroactive, photoactive and ion-exchange polymers. These polymers can be synthesized on a variety of microporous support membranes. To date, we have used microporous alumina (Anopore), a fluoropolymer (Gore-tex) and polycarbonate (Nucleopore) membranes.

* To whom correspondence should be addressed.

TABLE 1 Representative polymer films prepared using the photopolymerization method

Polymer number	Formula	Support
I	$\begin{array}{c} \text{---}(\text{CH}_2\text{---CH})_x\text{---} \\ \\ \text{C}\equiv\text{N} \end{array}$	Anopore
II	$\begin{array}{c} \text{CH}_3 \\ \\ \text{---}(\text{CH}_2\text{---C})_x\text{---} \\ \\ \text{C---O---CH}_3 \\ \\ \text{O} \end{array}$	Anopore
III	$\begin{array}{c} \text{Cl} \\ \\ \text{---}(\text{CH}_2\text{---C})_x\text{---} \\ \\ \text{Cl} \end{array}$	Anopore
IV	$\begin{array}{c} \text{---}(\text{CH}_2\text{---CH})_x\text{---} \quad (\text{CH}_2\text{---CH})_y\text{---} \quad (\text{CH}_2\text{---CH})_z\text{---} \\ \quad \quad \quad \quad \quad \quad \\ \text{C}_6\text{H}_5 \quad \text{C}_6\text{H}_5 \quad \text{C}_6\text{H}_4\text{SO}_3^-\text{Na}^+ \\ \quad \quad \quad \quad \quad \quad \\ \text{---}(\text{CH}_2\text{---CH})\text{---} \quad \text{CH}_2\text{---CH}_3 \quad \text{SO}_3^-\text{Na}^+ \end{array}$	Anopore, Nucleopore, Gore-tex
V	$\begin{array}{c} \text{---}(\text{CH}_2\text{---CH})_x\text{---} \quad (\text{CH}_2\text{---CH})_y\text{---} \quad (\text{CH}_2\text{---CH})_z\text{---} \\ \quad \quad \quad \quad \quad \quad \\ \text{C}_6\text{H}_5 \quad \text{C}_6\text{H}_5 \quad \text{C}_6\text{H}_4\text{Fe} \\ \quad \quad \quad \quad \quad \quad \\ \text{---}(\text{CH}_2\text{---CH})\text{---} \quad \text{CH}_2\text{---CH}_3 \quad \text{C}_6\text{H}_5 \end{array}$	Anopore

Polymer film I was polymerized from neat acrylonitrile; II, from neat methyl methacrylate; III, from neat vinylidene chloride; IV, from a solution 7.9 wt% sodium styrene sulphonate and 30 wt% technical-grade divinylbenzene in dimethylsulphoxide (62.1 wt%); V, from a 20 wt% vinylferrocene, 80 wt% technical-grade divinylbenzene mixture.

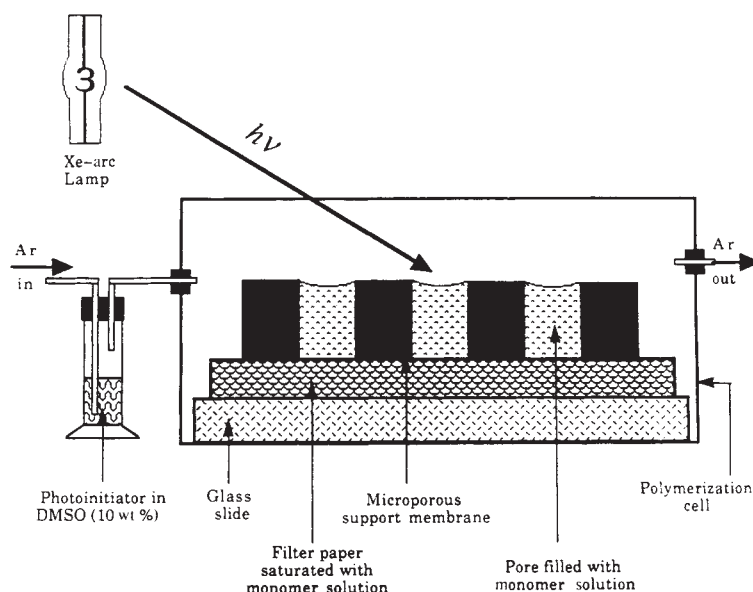
The mechanism of film formation depends on the nature of the polymer. For the cross-linked polymers (Table 1), the classical gel-point mechanism¹⁴ seems to operate. A series of electron micrographs taken during film formation are shown in Fig. 2. The total irradiation time was 30 minutes, but no visible evidence for film formation can be seen until about 20 minutes. This long induction period is characteristic of the gel-point mechanism¹⁴; furthermore, this mechanism is unique to cross-linked polymers.

Several other points are worth making. First, the incident ultraviolet light strikes the membrane surface at an acute angle

(Fig. 1). This minimizes the penetration depth of the photons and thus confines polymerization predominantly to the membrane surface (see Fig. 2). Nevertheless, at long polymerization times, polymer does propagate into the pores, and polymeric fibers are obtained. Second, a 30-min post-polymerization photocure was typically used to promote chain extension and further cross-linking. Finally, it would be possible to decrease the polymerization and photocuring times by using a high-intensity source (such as a laser)¹⁵.

Figure 2d is a scanning electron micrograph of a cross-section of a typical composite membrane. The film is so thin that it is

FIG. 1 Simplified schematic representation of photopolymerization procedure. DMSO, dimethylsulphoxide. The membrane was rotated during polymerization to ensure uniform photon fluxes. Excess monomer was removed, after photopolymerization, by aspiration.



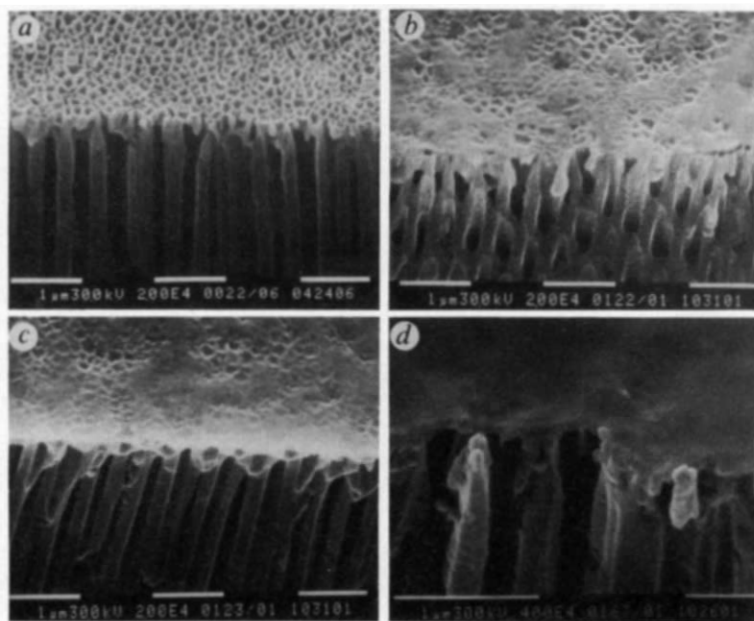


FIG. 2 Scanning electron micrographs of cross-sections and surfaces of a microporous support membrane (Anopore alumina filter) after 15 min (a), 20 min (b), 25 min (c) and 30 min (d) of photopolymerization. The polymer is polymer IV, Table 1. Scale bars are 1 μm . The membranes were coated with 25 nm of gold/palladium alloy before the micrographs were taken.

close to the resolution limit for this microscope (Phillips 505). The film thickness was determined by photographically enlarging such micrographs and making careful measurements of the film edge³. A film thickness of 40 nm was obtained. The gas-transport data, discussed below, provide independent corroboration for this measured film thickness.

The critical question is, of course, whether such extraordinarily thin films (Fig. 2d) are defect-free. We have used gas-transport measurements to address this question. We measured permeability coefficients for O₂ and N₂ in composite membranes based on polymer IV (Table 1) using the conventional single-gas permeation method^{16,17}, and took their ratios to obtain the O₂/N₂ selectivity coefficient ($\alpha_{\text{O}_2/\text{N}_2}$). Henis² has shown that if even a minute number of defects are present in such thin polymer films, the selectivity coefficient will assume the Knudsen value (in this case 0.9). In contrast, defect-free styrenics¹⁷ typically show $\alpha_{\text{O}_2/\text{N}_2}$ values of 4 to 6.

Selectivity coefficients of 8.0 were obtained for films of polymer IV which were as thin as 40 nm (Table 2). These data show that these ultrathin films are essentially defect-free². Furthermore, because film thickness enters into the calculation of permeability coefficient^{16,17}, these data corroborate the film thickness determined by electron microscopy.

Finally, the permeability and selectivity coefficients obtained for polymer IV are fairly high¹⁷. This suggests that these materials might be useful in commercial gas-separation systems. It is interesting to note that similar high selectivities and permeabilities have been measured for other highly ionic polymers^{4,5}. We are currently exploring the chemistry behind these

exceptional gas-transport properties. We are also exploring separations, sensor and drug-delivery applications of these new composite membranes. □

Received 5 February; accepted 30 April 1991.

1. Abelson, P. H. *Science* **244**, 1421 (1989).
2. Henis, J. M. S. & Tripodi, M. K. *Science* **220**, 11–17 (1983).
3. Liu, C., Espenscheid, M. W., Chen, W.-J. & Martin, C. R. *J. Am. chem. Soc.* **112**, 2458–2549 (1990).
4. Liang, W. & Martin, C. R. *Chem. Mater.* **3**, 390–391 (1991).
5. Anderson, M. R., Mattes, B. R., Reiss, H. & Kaner, R. B. *Science* **249**, 1412–1415 (1991).
6. Haggin, J. *Chem. Engng News* **68** (October 1), 22–26 (1990).
7. US Dept Energy Rep. No. DE90-011770 (US Department of Energy, Washington DC, 1990).
8. Vieth, W. R. *Membrane Systems: Analysis and Design* (Hanser, New York, 1988).
9. Martin, C. R. *et al. Polym. Mater. Sci. Engng* **64**, 286–287 (1991).
10. Dyer, C. K. *Nature* **343**, 547–548 (1990).
11. Sailor, M. J. *et al. Science* **249**, 1146–1149 (1990).
12. Tierney, M. J. & Martin, C. R. *J. electrochem. Soc.* **137**, 2005–2006 (1990).
13. Roffey, C. G. *Photopolymerization of Surface Coatings*, 80–83 (Wiley, New York, 1982).
14. Williams, D. J. *Polymer Science and Engineering*, 67 (Prentice-Hall, Englewood Cliffs, New Jersey, 1971).
15. Billmeyer, F. W. Jr *Textbook of Polymer Science*, 65–67 (Wiley, New York, 1984).
16. Koros, W. J., Chan, A. H. & Paul, D. R. *J. Membrane Sci.* **2**, 165–190 (1977).
17. Puleo, A. C., Muruganandam, N. & Paul, D. R. *J. Polym. Sci. B* **27**, 2385–2406 (1989).

ACKNOWLEDGEMENTS. This work was supported by the Air Force Office of Scientific Research.

Self-assembling phospholipid filaments

A. S. Rudolph, B. R. Ratna* & B. Kahn†

Center for Bio/Molecular Science and Engineering, Code 6090, Naval Research Laboratory, Washington DC 20375–5000, USA

* Department of Biochemistry, Georgetown University Medical Center, Washington DC 20007, USA

† Geo-Centers Inc, 10904 USA Indian Head Highway, Fort Washington, Maryland 20744, USA

AQUEOUS dispersions of double-chain phospholipids spontaneously assemble into closed bilayers called vesicles (or liposomes). Although the vesicles are in general topologically spherical, cylindrical¹ and helical² liposomes have sometimes been observed. We present here video-enhanced microscopic studies of a diacetylenic phospholipid dispersed in ethanol/water, which reveal the existence of unusual bilayer morphologies. On cooling the dispersion from the isotropic phase, we have observed the formation of long (of the order of hundreds of micrometres), thin (0.2–2 μm) filaments, which fluctuate strongly. When the tem-

TABLE 2 Gas-transport properties of composite membranes based on polymer IV*

Film thickness† (nm)	Permeability coefficient‡		Separation factor O ₂ /N ₂
	O ₂	N ₂	
2,300	0.54	0.066	8.2
150	0.55	0.068	8.0
40	0.57	0.071	8.0

* See Table 1.

† Determined from electron micrographs. Irradiation times were 180 min (2.3 μm), 60 min (150 nm) and 30 min (40 nm).

‡ Calculated from gas flux data and film thickness. Units are (10^{–10} cm³ (STP) cm)/(cm² s cm(Hg)).

perature is decreased further, the filaments rapidly retract into a mass of lipid. At constant temperature, on the other hand, the filaments transform into torus or ring-like vesicles. Such non-spherical structures have been predicted theoretically^{3,4} but not previously observed experimentally.

The double-chain phospholipid used in our experimental study was 1,2-bis(10,12-tricosadiynoyl)-sn-glycero-3 phosphocholine^{5,7}, a lecithin with polymerizable diacetylenic groups midway in each of the otherwise saturated acyl chains. Video-enhanced microscopic observations were made on the lipid dispersed in a solvent mixture of 70/30 by volume of ethanol/water and taken in flat glass capillaries (0.05–0.2-mm path length). The phospholipid concentrations studied were 5–20 mg per ml. Calorimetric studies showed that the lipid in the solvent mixture has a chain melting temperature T_m of 36 °C. The corresponding crystallization temperature on cooling is 33 °C. The sample, under a polarizing microscope, shows the typical 'oily streak' texture of the L_α phase at temperatures above T_m . As the temperature is increased in the L_α phase, the lipid spreads and appears as thin plates until dissolution into the solvent is observed at 55–60 °C, leaving an optically clear solution. Quasi-elastic light-scattering measurements at these temperatures indicate that it may be an isotropic solution of lipid molecules with no self association.

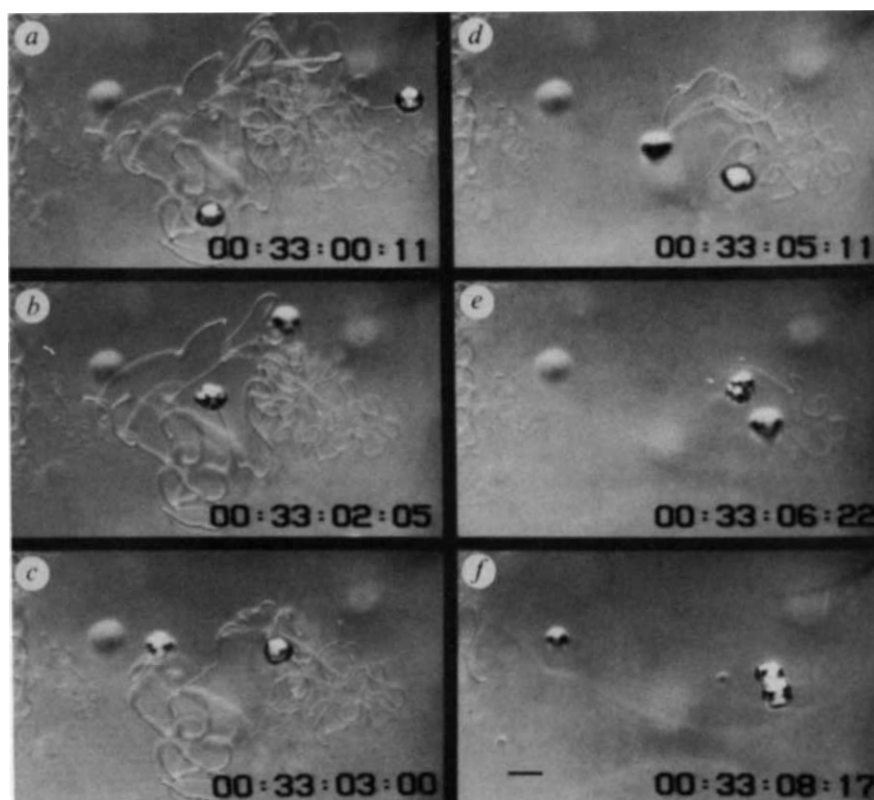
On cooling the isotropic solution under the microscope, we first observed long, thin filaments at ~49 °C. Considering the geometric shape of the lipid⁸ as well as the fact that the sample showed a lamellar L_α phase when heated, we believe that these filaments are tubes made of fluid, liquid crystalline bilayers. These filaments often have a bulb-like appearance at their ends, from which the filaments are sometimes observed to grow. The thermal fluctuations of the filaments can be clearly seen under the microscope. They appear as an undulation when the direction of fluctuation is in the plane of focus and as a defocusing effect when the fluctuation direction is out of the focal plane. The strength of the fluctuations seems to be dependent on the filament diameter. Thinner filaments ($\leq 0.5 \mu\text{m}$) seem to undulate more rapidly and with a higher amplitude than thicker

filaments (1–2 μm). This may be due to the larger number of bilayers present in the thicker filaments, which is supported by the higher contrast shown by thicker filaments in the interference microscope.

The morphological changes in the filaments depend on whether the sample temperature is lowered continuously or is held constant. If the temperature is lowered towards the crystallization temperature, the filaments start to retract into the balls at their ends. The increasing mass of lipid at the tip moves around engulfing the filaments. This sequence of events is shown in Fig. 1a–f. If we increase the lipid concentration, the density of the filaments increases, and often several filaments retract into the same lipid mass (Fig. 2a–f). The filaments exist over a wide range of temperature (49 °C–33 °C) and the temperature at which the retraction occurs is not well defined. At any given temperature the filaments coexist with the retracted lipid masses, the density of the former decreasing and the latter increasing as the temperature is lowered. The resolution of the microscope is insufficient to determine whether the filaments in the collapsed structure remain as filaments or fuse and assume another morphology. We are making efforts to freeze-fracture the collapsed lipid mass and observe its structure in a transmission electron microscope.

If the temperature is held constant, the filaments undergo a different morphological change. The segments of the filaments that are in close proximity fuse along their length creating thicker filaments (Fig. 3a–d). The thermodynamic driving force for the sudden collapse as well as the formation of thicker filaments is probably the membrane-membrane binding. At larger separations the interaction between the filaments is probably dominated by the long-range repulsion due to fluctuation forces⁹. But if the filaments approach each other closely enough for the van der Waals attraction to become stronger than the entropic repulsion¹⁰, the filaments bind to each other. Once the filaments are attached at a point, the 'zipper-closing' effect (Fig. 3) occurs, resulting in a thicker filament. If we continue to hold the temperature constant, the filament often coils up to form a large torus-shaped vesicle (Fig. 4a–f). These toroidal structures are

FIG. 1 Representative video images of a single long filament during retraction. a, Typical single filament formed on cooling (cooling rate is 1 °C per min) from the isotropic phase. Temperature is 40.4 °C. The two ends of the filaments are attached to balls of lipid. As the temperature is lowered (a–f), the two ends move along the length of the filament, engulfing them in the process. The phenomenon occurs very rapidly, the whole sequence takes ~8 s (time-code is shown at the bottom of each photograph). Scale bar, 20 μm .



thermally stable and long-lived. This equilibrium shape corresponds to a state of minimum elastic free energy.

The elastic energy per unit area can be written as¹¹

$$g_c = 1/2 k_c (c_1 + c_2 - c_0)^2$$

where c_1 and c_2 are the principal curvatures, c_0 is the spontaneous curvature and k_c is the bend elastic modulus. For a flaccid vesicle, the bilayer is practically free of lateral tension and the shape is solely determined by the curvature elasticity¹². Depend-

ing on the value of c_0 , the shape can vary from oblate to prolate spheroids, and further into cylinders and long tubes. Recently, a more generalized free-energy equation^{3,4} has been developed, taking into consideration not only the effect of the bending modulus but also the osmotic pressure difference between the outside and inside of the vesicle, and the tensile strength. The important outcome of this theory is the existence of anchor-ring vesicles³ and very long cylindrical filaments⁴. Though the anchor-ring vesicle dictates the ratio of the generating radii to

FIG. 2 Representative video images of a very high density of filaments formed at higher lipid concentrations (~20 mg per ml). Cooling rate is 1 °C per min; temperature is 37.8 °C. In this example, multiple filaments seem to end on a single lipid mass. In such a case the filaments retract into the fairly stationary lipid mass. Scale bar, 20 µm. The 'staircase' steps seen on the filaments are video artefacts.

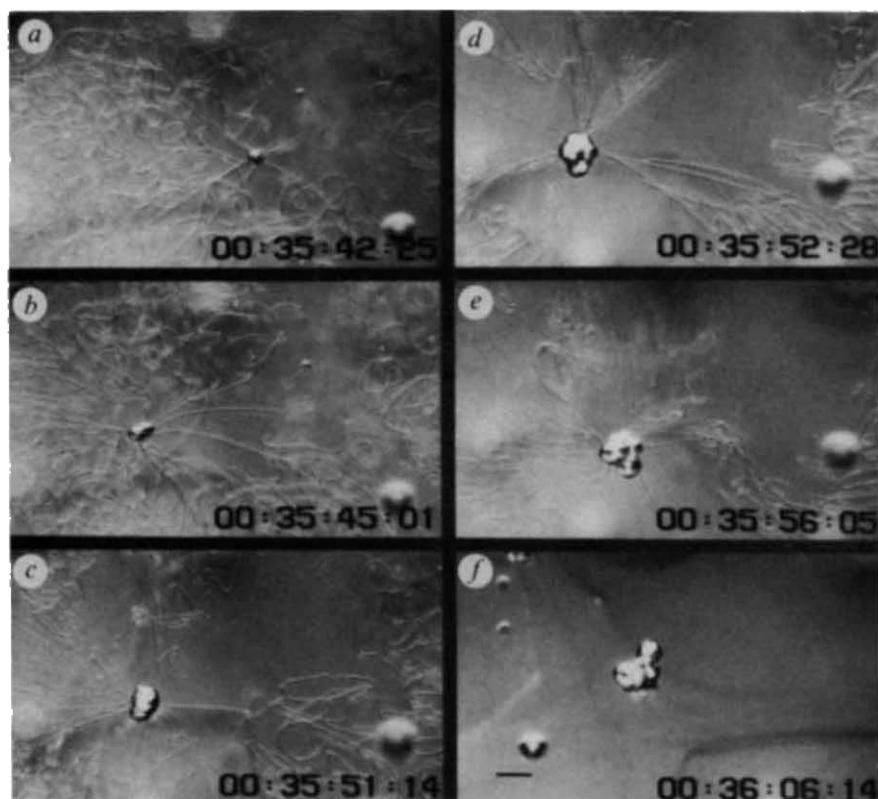


FIG. 3 Representative video images of the formation of thicker filaments at constant temperature of 35.6 °C. The arrow in *a* shows the beginning of the fusion of thinner filaments. As the thinner filaments are drawn together, the length of the resulting thicker one increases (follow the arrow from photographs *a-d*). The elapsed time between *a* and *d* is ~1.5 min. Scale bar, 10 µm.

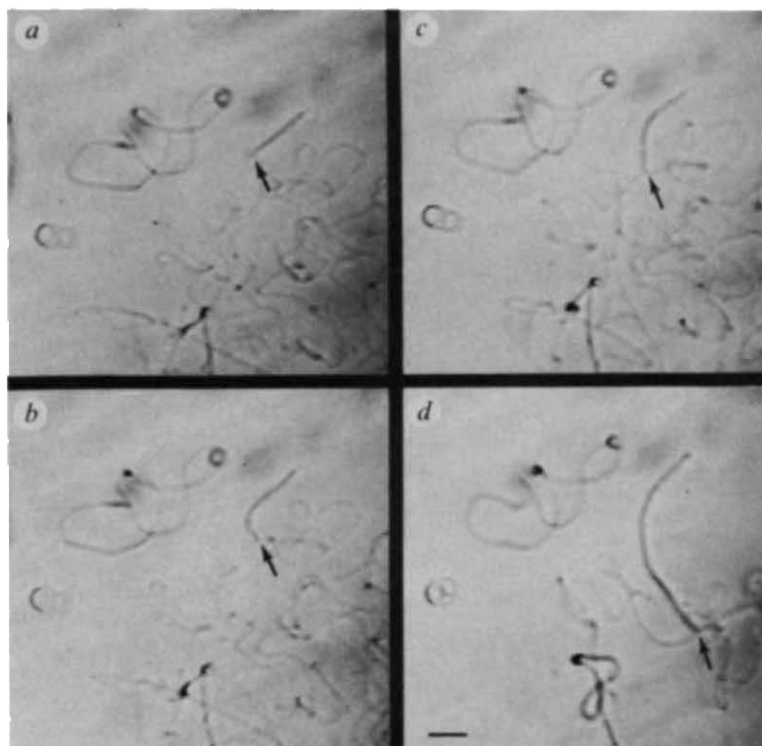
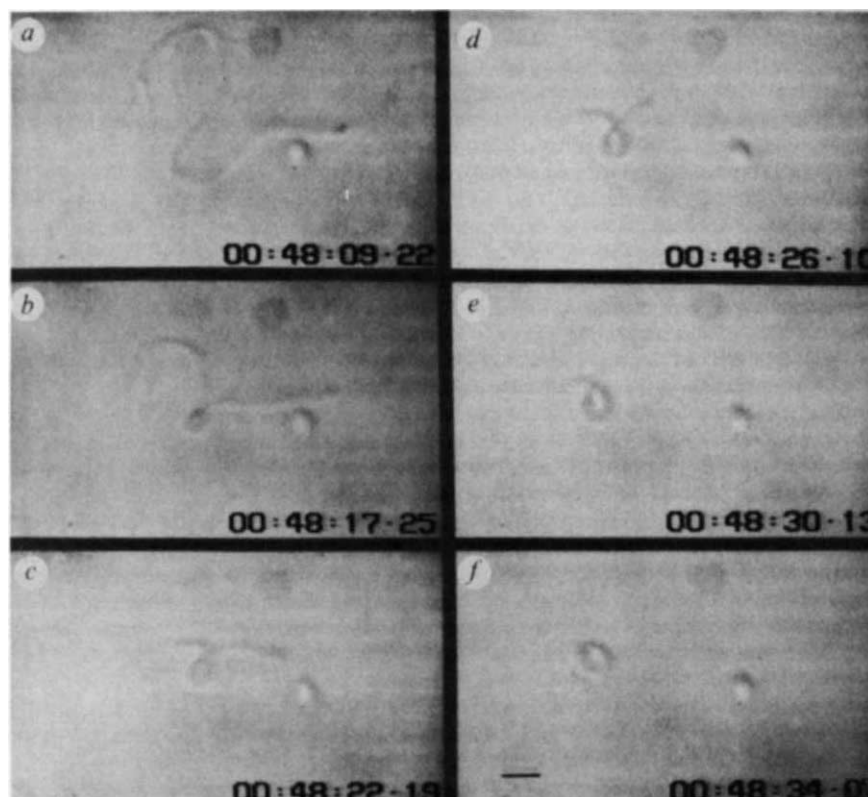


FIG. 4 Formation of the torus from a filament. Temperature is 46 °C. *a*, Thicker filament formed after the fusion of thinner filaments. *b*, If the temperature is still held constant, an initial loop forms. *c–f*, As time elapses, the rest of the filament is drawn into the loop to form a torus (*f*). Scale bar, 10 μm .



be $1/\sqrt{2}$, this is only the upper limit, and the theory does not exclude the formation of larger tori with smaller radius ratio, like the ones observed here.

Although assemblies of thin filaments in synthetic lipids are unusual, a variety of lipid-based, nonspherical shapes are observed in biological systems. Of particular interest are long, thin filaments that have been observed in erythrocyte membranes subjected to ageing or other pathological conditions¹³. But one has to be extremely cautious in comparing a structure obtained from a synthetic lipid to a biological membrane containing several lipid species. Although the reasons for the heterogeneity of lipids in biomembranes are not clear, one factor to be considered^{14–16} is the different packing configurations that may be necessary to stabilize regions of different spontaneous curvature needed to maintain the shape of the cell. It is important to mention here that 1,2-di(tricosanoyl)-sn-glycero-3 phosphocholine, a saturated analogue of the diacetylenic lipid, observed under the same solvent conditions, did not show any unusual morphologies. Clearly, the diacetylenic group in the middle of the acyl chain strongly influences the spontaneous curvature. More experiments are needed to understand the part played by the diacetylenic moiety, which is rather rigid compared with the unsaturated groups generally found in lipids of biological origin. □

Received 25 March; accepted 24 May 1991.

- Schneider, M. B., Jenkins, J. T. & Webb, W. W. *Biophys. J.* **45**, 891–899 (1984).
- Lin, K.-C., Weis, R. M. & McConnell, H. M. *Nature* **296**, 164–165 (1982).
- Ou-Yang Zhong-can *Phys. Rev. A* **41**, 4517–4520 (1990).
- Ou-Yang Zhong-can & Helfrich, W. *Phys. Rev. A* **39**, 5280–5288 (1989).
- Singh, A. & Schnur, J. M. *Synth. Commun.* **16**, 847–852 (1986).
- Chapman, D. US Patent No. 4, 348, 329 (1982).
- Hub, H.-H., Hupfer, B., Koch, H. & Ringsdorf, R. *Angew. Chem. Int. Ed. Engl.* **19**, 938–940 (1980).
- Israelachvili, J. N. *Intermolecular and Surface Forces* (Academic, London, 1985).
- Evans, E. A. & Parsegian, V. A. *Proc. natn. Acad. Sci. U.S.A.* **83**, 7132–7136 (1986).
- Cevc, G. & Marsh, D. *Phospholipid Bilayers—Physical Principles and Models* (Wiley, New York, 1987).
- Deuling, H. J. & Helfrich, W. *J. Phys.* **37**, 1335–1345 (1976).
- Deuling, H. J. & Helfrich, W. *Blood Cells* **3**, 713–720 (1977).
- Bessis, M. *Living Blood Cells and their Ultrastructure*, 61 and 170 (Springer, New York, 1973).
- Cullis, P. R., Hope, M. J. & Tilcock, C. P. S. *Chem. Phys. Lipids* **40**, 127–131 (1986).
- Cullis, P. R., Hope, M. J., de Kruijff, B., Verkleij, A. J. & Tilcock, C. P. S. in *Phospholipids and Cellular Regulations* Vol. 1 (ed. Kuo, J. F.) (CRC, Boca Raton, Florida, 1985).
- de Kruijff, B. *Nature* **329**, 587–588 (1987).

ACKNOWLEDGEMENTS. This research is supported by the Office of Naval Research.

Role of eddy pumping in enhancing primary production in the ocean

Paul G. Falkowski,* David Ziemann,† Zbigniew Kolber* & Paul K. Bienfang†

* Oceanographic and Atmospheric Sciences Division, Brookhaven National Laboratory, Upton, New York 11973, USA

† Oceanic Institute, Makapuu Point, Waimanalo, Hawaii 96795, USA

IN steady-state models of primary production in the open ocean, the upward fluxes of nutrients are balanced by the export of particulate production to the ocean depths. Export production represents the biological effect of carbon production on the net exchange of carbon dioxide between the atmosphere and the ocean. Geochemical estimates of exported production, based on calculations of rates of oxygen usage¹ or heat fluxes² are two to three times as high as those determined from biological measurements^{3–5}. One possible explanation for the differing estimates is that export production, calculated from independent geochemical signals, is too high. Another is that biological measurements severely undersample episodic nutrient injections into the photic zone^{1,4}. Eddy pumping represents one of the possible mechanisms of nutrient injection¹. Here we examine the enhancement of production by a cyclonic eddy in the subtropical Pacific with instrumentation that allows us to overcome the sampling problem. Our results reveal that eddy pumping markedly stimulates primary production, and that phytoplankton in the upper oligotrophic ocean outside the eddy are not growing near their maximum relative specific growth rates. But if the relative enhancement of production is typical, our results suggest that eddy pumping would enhance total primary production by only ~20%.

Mesoscale cyclonic eddies occur frequently in the lee of the island of Hawaii^{6,7}. These eddies, which may range in diameter

from 50 to 100 km, are produced as a result of wind stress circulation patterns⁶ and are consistently found throughout most of the year⁷. The cyclonic rotation leads to a doming or localized upwelling of deep, nutrient-rich waters into the euphotic zone (Fig. 1). Individual eddies have a life span of 6–12 weeks, before the momentum dissipates and the doming relaxes. The fairly predictable occurrence of eddies off Hawaii, coupled with recent advances in instrumentation⁸, allowed us to evaluate the response of phytoplankton to the physical dynamics of such eddies.

Sections of temperature through an eddy, sampled along a track of 156° 22' W in September 1989, reveal a doming of isotherms in the upper 200 m (Fig. 2a), which is typical of a cyclonic eddy in the Northern Hemisphere. The distribution of nitrate demonstrates the effect of eddy energy in pumping nutrient-rich water into the surface layer (Fig. 2b). The isopleths of nitrate closely follow the isotherms; slight deviations seem to be artefacts of contouring. The nitrate contour at 2 μ M, for example, is elevated from deeper than 110 m (the depth at which it is generally found in Hawaiian waters⁹) to less than 75 m. Nitrate concentrations at the base of the euphotic zone (~100 m) increased from 2.5 μ M to nearly 4 μ M at the centre of the eddy.

Evidence that nutrient pumping by the eddy affects phytoplankton abundance is seen in the distribution of chlorophyll (Fig. 2c). The central subtropical Pacific Ocean is typified by a deep chlorophyll maximum layer of concentration 0.2–0.4 μ g per litre, generally located at ~90 m (ref. 9). This pattern was typical for stations located outside the eddy, generally showing a chlorophyll maximum layer at 90 m with levels approaching 0.3 μ g per litre. At the centre of the eddy, however, the chlorophyll maximum was at 30 m, with a concentration of slightly less than 0.3 μ g per litre.

The physiological response of the phytoplankton to the eddy pumping was shown by a marked increase in the level of the maximum change in variable chlorophyll fluorescence, $\Delta\phi_{\text{sat}}$. This parameter, measured with a custom-built pump-and-probe fluorimeter⁸, is related to the maximum efficiency of conversion of energy by photosynthesis. The spatial patterns of $\Delta\phi_{\text{sat}}$ are independent of growth irradiance but related to the supply of nutrients^{8,10}; higher values of $\Delta\phi_{\text{sat}}$ are associated with increased nutrient supply⁸. The molecular basis of the changes in $\Delta\phi_{\text{sat}}$ seems to be the effect of limitation by nutrients (especially inorganic nitrogen) on the transfer of excitation energy from the light-harvesting antenna complexes to photosystem II reaction centres¹⁰. Low values of $\Delta\phi_{\text{sat}}$ were found both below the 100-m isobath, where cells become increasingly isolated from the upper mixed layer, and in the upper 50 m outside the eddy

TABLE 1 Productivity and growth rates at stations within and outside the eddies

Station	1% light depth (m)	$\bar{\mu}$	Integrated chlorophyll (mg m ⁻²)	Integrated primary productivity (mg C m ⁻² d ⁻¹)
6	96	0.140	8.51	173.1
7	148	0.143	21.25	492.7
8	96	0.192	10.3	201.31
9	94	0.172	9.68	216.12
10	86	0.357	20.48	657.26
11	100	0.084	19.56	244.13
3	88	0.168	5.27	117.86

Depth at which light is reduced to 1%, mean relative growth rates ($\bar{\mu}$), averaged for the euphotic zone, integrated chlorophyll, and primary productivity integrated from the surface to 1% light depth. Primary productivity was calculated from profiles of variable fluorescence as described in ref. 15 and Fig. 3.

vortex. The largest variable fluorescence signal was observed at station 10. At station 4, another station in the eddy, $\Delta\phi_{\text{sat}}$ was markedly increased in the upper 50 m (data not shown). Variable fluorescence outside the eddy (for example at station 6) was highest at the depth of the nitricline.

In nitrogen-limited phytoplankton, $\Delta\phi_{\text{sat}}$ is a hyperbolic function of maximum relative growth rate¹⁰. The variability in the fluorescence data shows that phytoplankton growth rates are nutrient-limited in the upper ocean outside the eddy. From the hyperbolic relationship between $\Delta\phi_{\text{sat}}$ and relative specific growth rates, we calculated the enhancement of the relative maximum growth rate of chlorophyll resulting from eddy pumping of nutrients into the euphotic zone (Table 1). These results strongly suggest that the regeneration of nutrients by heterotrophic grazers is not sufficient to maintain the phytoplankton near their maximum relative specific growth rates¹¹. Consequently, we conclude that in the upper oligotrophic Pacific Ocean, the rate of phytoplankton growth is physiologically limited by the availability of nutrients.

We assessed the effect of eddy pumping on integrated water-column primary productivity (Table 1). Continuous vertical profiles of primary productivity were made at each station during daylight hours using a custom-built pump-and-probe fluorimeter^{10,12} interfaced to a conductivity-temperature-depth instrument. From measurements of $\Delta\phi$ under ambient irradiance and in the dark, we calculated the contribution of photochemical and non-photochemical processes to the overall fluorescence yield¹³ and used the equations of Genty *et al.*¹⁴ to derive the instantaneous rates of photosynthetic electron transport. The fluorescence-based productivity rates have been previously compared with discretely sampled ¹⁴C measurements in the North Atlantic¹⁵; the results indicate that the fluorescence-based estimates of production, using the pump-and-probe fluorimeter, are significantly linearly correlated with radiocarbon measurements. Using derived parameters from that analysis for our specific data set, we further compared discrete radiocarbon production measurements with estimates of photosynthetic electron transport calculated from the changes in the quantum yields of fluorescence, the incident irradiance and the measured absorption cross-sections¹⁵ of photosystem II (Fig. 3). Our results indicate that estimates of primary production based on the *in situ* pump-and-probe technique are robust and comparable in precision to radiocarbon estimates of production.

The fluorescence-based estimates of production showed that the cyclonic eddy increased the vertically integrated primary production to 3.5 times that of the adjacent areas (Table 1). This enhancement of production was accompanied by a 67% increase in the integrated chlorophyll-specific rate of photosynthesis.

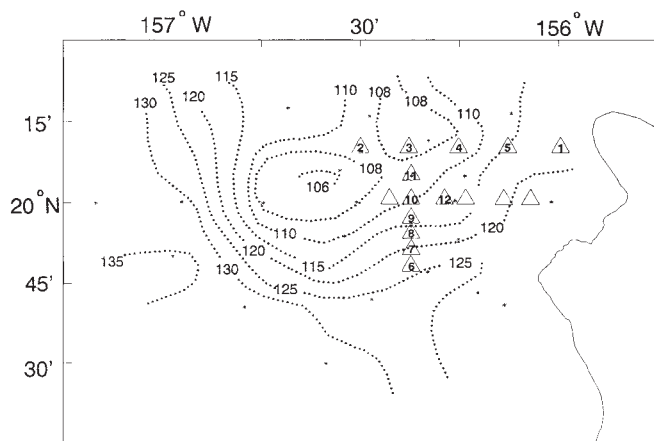


FIG. 1 Contour plot of depths of the isotherm at 23°C from an areal expendable bathythermograph survey taken on 29 August 1989, showing the position of a mesoscale eddy, centred at 156° 30' W and 20° 05' N, off the island of Hawaii. Station numbers occupied by ship nine days later are indicated by triangles.

FIG. 2 *a*, Vertical section along 156° 22' W taken 8–10 September 1989 (see Fig. 1), showing the doming of isotherms and the thermal eddy signature, centred at station 10. *b*, Nitrate concentrations were measured on frozen samples by an AutoAnalyzer at the University of Hawaii. *c*, Section showing chlorophyll *a* concentrations. Samples were filtered on Whatman GF/F glass-fibre-type filters, and chlorophyll was extracted in 90% aqueous acetone overnight at -20 °C. Chlorophyll was determined by fluorescence before and after acidification. *d*, Section of the maximal change in variable fluorescence ($\Delta\phi_{\text{sat}}$) measured with a pump-and-probe fluorimeter⁸ with an 80- μs delay between the pump and probe flash. The maximum observed value of $\Delta\phi_{\text{sat}}$ is 1.27. The spatial variability in $\Delta\phi_{\text{sat}}$ reflects the spatial variability in the relative specific growth rates of phytoplankton.¹⁰

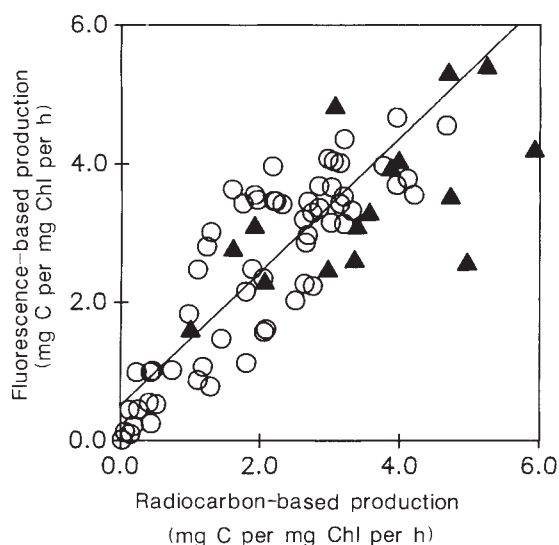
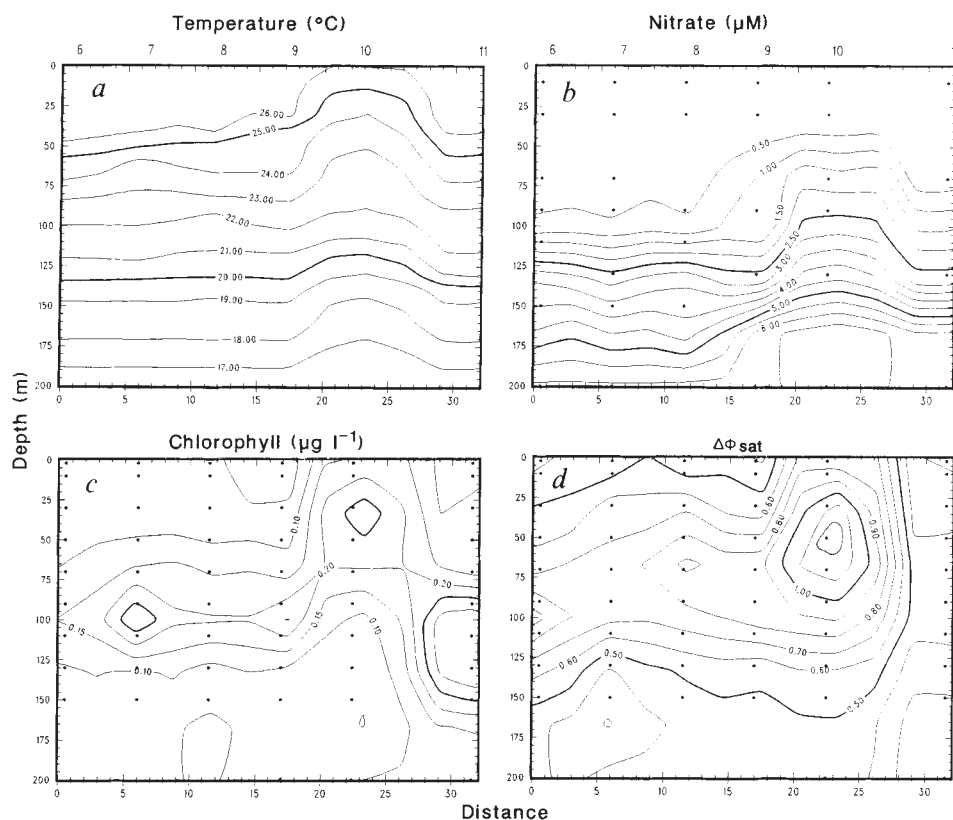


FIG. 3 Comparison between fluorescence-based estimates of *in situ* photosynthesis using the pump-and-probe technique and radiocarbon-based measurements (Chl stands for chlorophyll). The fluorescence-based estimates were calculated from incident irradiance, the absorption cross-section of photosystem II, the change in the quantum yield of fluorescence and the turnover time of the photosynthetic electron transport chain.¹⁵ Photosynthesis was calculated from the equation $P = I_z \sigma_{\text{PSII}} q \phi_e n \Delta\phi_{\text{sat}}$, where I_z is incident photosynthetically active radiation (400–700 nm) at depth z , σ_{PSII} is the absorption cross section of photosystem II,¹⁵ $q\phi_e$ is non-photochemical quenching,¹⁴ ϕ_e is the yield of electron transfer, n is the concentration of photosystem II reaction centres and $\Delta\phi_{\text{sat}}$ is the maximum variable fluorescence yield. Radiocarbon-based measurements were obtained simultaneously from 2–6-hour incubations at irradiance levels simulating the depth at which the water samples were collected. O, data obtained in the northwest Atlantic Ocean in 1988 and 1989; ▲, data obtained in the subtropical Pacific in September 1989. The line is the best fit to a linear regression (sample number $n=76$; correlation coefficient $r=0.82$).

These results demonstrate that eddy pumping markedly affects local phytoplankton production in the subtropical ocean. The systematic spatial changes in $\Delta\phi_{\text{sat}}$ strongly imply that the rate of regeneration of nutrients in the upper ocean by heterotrophic processes, contrary to expectations¹¹, is not sufficient to sustain phytoplankton at their maximum relative growth rates.

Is the enhancement of primary production by eddy pumping sufficient to resolve the discrepancy between geochemical and biological estimates of new production? If the empirical relationships between new and total production obtained from both biological rate measurements³ and sediment trap studies¹⁶ are valid, the differences between exported carbon in a cyclonic eddy and outside the eddy vary by a factor of ~ 3.5 . Wyrki *et al.*¹⁷ estimated that the ratio between mean and eddy energy in central ocean gyres is ~ 0.05 . If our observations are typical of the enhancement of production caused by eddy pumping, then eddy pumping would enhance total production by $\sim 20\%$ above the mean in the central ocean gyres, and its effect on export production would be less than 20%. We suggest that if eddy pumping were to account for the discrepancy between biological and geochemical estimates of export production, then the enhancement of primary production by eddy pumping would have to be about an order of magnitude greater than we have reported. Although that possibility cannot be ruled out, it seems likely that the signal would be readily detectable in the satellite colour images of the central ocean gyres (Fig. 2c). So far, mesoscale variability of that magnitude has not been reported from satellite data analyses of the oligotrophic ocean. □

Received 21 February; accepted 21 May 1991.

1. Jenkins, W. J. *Nature* **331**, 521–523 (1988).
2. Hayward, T. L. *Deep-Sea Res.* **34A**, 1593–1627 (1987).
3. Eppley, R. W. & Peterson, B. J. *Nature* **282**, 677–680 (1979).
4. Platt, T. & Harrison, W. G. *Nature* **318**, 55–58 (1985).
5. Chavez, F. P. & Barber, R. T. *Deep-Sea Res.* **34**, 1229–1243 (1987).

6. Patzert, W. C. *Tech. Rep. No. 69-8* (Hawaii Institute of Geophysics, University of Hawaii, 1969).
7. Lobel, P. S. & Robinson, A. R. *Deep-Sea Res.* **33**, 483-500 (1986).
8. Kolber, Z., Wyman, K. D. & Falkowski, P. G. *Limnol. Oceanogr.* **35**, 72-79 (1990).
9. Bienfang, P. K. & Szyper, J. P. *Deep-Sea Res.* **28**, 981-1000 (1981).
10. Kolber, Z., Zehr, J. & Falkowski, P. *Plant Physiol.* **88**, 923-929 (1988).
11. Laws, E. A., DiTullio, G. R. & Redlge, D. G. *Limnol. Oceanogr.* **32**, 905-918 (1987).
12. Falkowski, P. G., Wyman, K. D., Ley, A. C. & Mauzerall, D. C. *Biochim. biophys. Acta* **849**, 183-192 (1986).
13. Weiss, E. & Berry, J. A. *Biochim. biophys. Acta* **894**, 189-208 (1987).
14. Genty, B., Briantais, J. M., and Baker, N. R. *Biochim. biophys. Acta* **932**, 102-115 (1989).
15. Falkowski, P. G. & Kolber, Z. *Current Research in Photosynthesis IV* (ed. Baltscheffsky, M.) 923-926 (Kluwer, London, 1990).
16. Betzer, O. R. et al. *Deep-Sea Res.* **31**, 1-72 (1984).
17. Wyrki, K., Magaard, L. & Hager, J. *J. geophys. Res.* **81**, 2641-2646 (1976).

ACKNOWLEDGEMENTS. We thank P. Minnett, K. Wyrki, W. Jenkins and D. Wallace for discussions and suggestions, and M. Oliazola, L. Conquest and C. Wilson for technical support. This research was supported by the NSF, the US Department of Energy and NASA.

Inverse square-root dependence of mid-ocean-ridge flank roughness on spreading rate

Alberto Malinverno

Lamont-Doherty Geological Observatory of Columbia University, Palisades, New York 10964, USA

THE topographic roughness of mid-ocean-ridge flanks is known to increase with decreasing spreading rate¹⁻⁴, but the exact form of this relationship has not been established. As the topographic features that make up ridge flank roughness (abyssal hills) are created near the axes of mid-ocean ridges by faulting and volcanism⁵⁻⁷, the relationship between roughness and spreading rate may shed light on the process of crustal accretion at spreading centres. Here I present measurements of ridge flank roughness on profiles crossing the world mid-ocean-ridge system, which show that roughness is proportional to the inverse square-root of the spreading rate. This result is consistent with some very simple inferences of how the topographic roughness of mid-ocean-ridge flanks scales with the lithospheric thickness near ridge axes, as determined by thermal models.

The data set used in this study (Fig. 1) consists of 101 profiles, which are 500-1,000 km long, have an average sampling interval of 2.5 km or less, are roughly oriented along flow lines and do not cross major fracture zones. This data set does not include profiles near known hot spots, or from areas where the ridge axis is anomalously shallow or deep (notable examples are the Reykjanes Ridge⁸ and the Australian-Antarctic Discordance⁹).

The measure of roughness used here is the root-mean-square (r.m.s.) roughness defined as the square-root of the average squared deviation about a linear trend. The statistical characteristics of ocean floor roughness are now well known¹⁰⁻¹⁴; particularly relevant here are the observations that r.m.s. roughness as measured on any profile increases with the length of the profile (for lengths up to a few tens of kilometres) and that the estimation of r.m.s. roughness can be severely biased by the presence of even a few large seamounts. Seamounts are volcanic features up to several thousands of metres high created both on and off the axes of mid-ocean ridges. As their connection to processes on the ridge axis is not well established, here it is desirable to measure ridge flank roughness excluding seamounts. The final estimate of r.m.s. roughness is taken as the median (rather than the arithmetic mean) of the r.m.s. roughnesses measured on several 100-km-long windows in the profile. This estimate of roughness is thus independent of the profile length and should not be biased by the presence of large seamounts (unless more than 50% of the windows contain seamounts).

I have compared the r.m.s. roughness measured for each profile with the present local spreading rate¹⁵ (an alternative approach is to consider the variation in roughness along flow lines where spreading rate changes significantly; D. E. Hayes and K. A. Kane, manuscript in preparation). When r.m.s. roughnesses are plotted linearly against spreading rates (Fig. 2a), it seems that there are two distinct regions, with roughness being strongly dependent on spreading rate for full rates <70 mm yr⁻¹, but becoming roughly constant at greater rates. Small, Sandwell and co-workers^{16,17} have proposed an analogous interpretation of the relationship between spreading rate and the amplitude of gravity anomalies on ridge axes and flanks, with an abrupt change at a full spreading rate of ~70 mm yr⁻¹. When the topographic roughness data are plotted in log-log coordinates, however, they follow a straight line with a slope of about -1/2 (the best-fit slope is -0.539 ± 0.036 ; Fig. 2b). This relationship corresponds to a power law, with r.m.s. roughness being inversely proportional to the square root of the spreading rate. In other words, the change in gradient apparent in Fig. 2a at 70 mm yr⁻¹ can be as well explained by a simple, continuous law relating ridge flank roughness to spreading rate, and the rate of 70 mm yr⁻¹ may have no particular significance. Note also that there is a considerable scatter around any smooth relationship between roughness and spreading rate, implying that there are additional factors besides spreading rate that affect ridge flank roughness (likely candidates may be local mantle temperature and distance from transform faults).

It has been suggested¹⁸⁻²¹ that the strength (thickness) of the

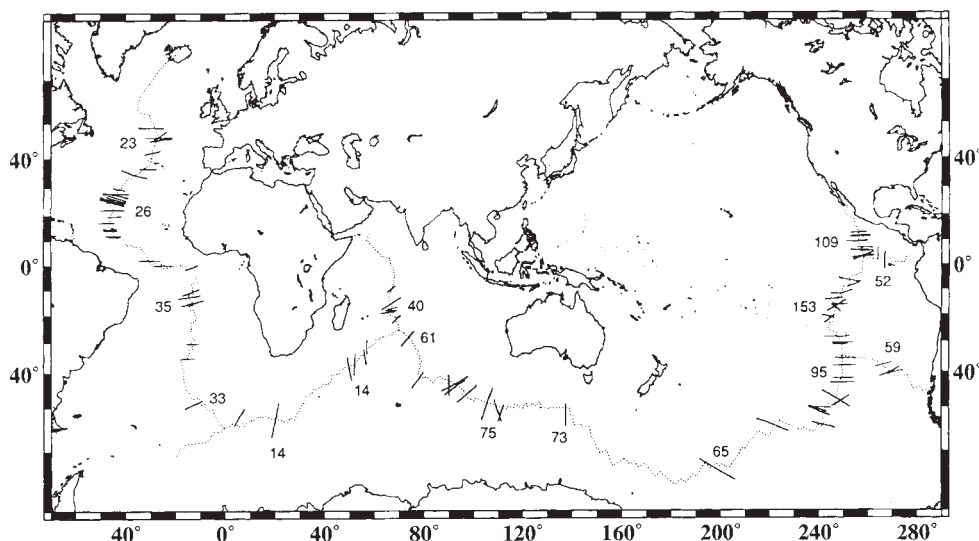


FIG. 1 The 101 profiles used in this study (continuous lines). The dotted line is the trace of the mid-ocean-ridge system; numbers are full spreading rates¹⁵.

lithosphere near ridge axes increases with decreasing spreading rate. A simple explanation for the observed relationship between spreading rate and ridge flank roughness is thus that the thickness of the lithosphere near the ridge axis controls the processes that create topographic roughness (volcanism and normal faulting). Models of the thermal structure of ocean lithosphere that incorporate the effect of mantle flow and account for hydrothermal circulation^{20,22} suggest that the lithospheric thickness at mid-ocean-ridge axes, measured as the depth of a given isotherm, is roughly proportional to the inverse of the spreading rate (Fig. 3). If lithospheric thickness is proportional to inverse spreading rate and topographic roughness is proportional to the inverse square root of the spreading rate, the roughness should be proportional to the square-root of lithospheric thickness.

To test this prediction, let us consider some simple models for the creation of topographic roughness by volcanism and faulting. Suppose ridge flank roughness is caused by the superposition of a number of volcanic constructions of average height h and width w , placed at discrete random locations (Fig. 4a, b). This procedure is equivalent to emplacing volcanoes at random times in a location fixed at the ridge axis, and moving this location in discrete steps at half the spreading rate. In this simple case, the heights will have a Poisson distribution, with a mean thickness for the volcanic layer $t = hNp$ and a r.m.s. roughness equal to $h(Np)^{1/2}$, where N is the total number of volcanic constructions, and p is the probability for a construction to fall in a given location (on a profile of length x , $p = w/x$). Combining these two expressions, topographic roughness can be expressed in terms of the thickness of the volcanic layer and is equal to $(th)^{1/2}$. Assuming that the thickness of the volcanic layer is constant, roughness must be proportional to the square root of the average height of volcanic constructions. This is obviously

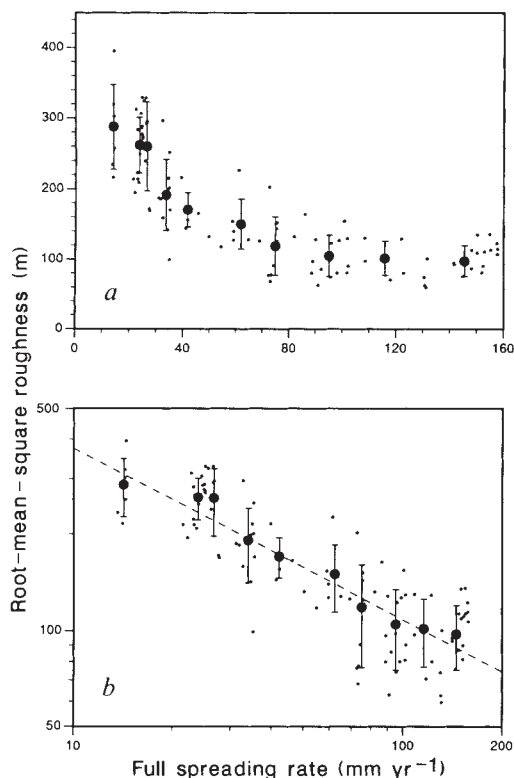


FIG. 2 Root-mean-square roughness of the topographic profiles of Fig. 1 plotted against local full spreading rates¹⁵ in *a*, linear and *b*, log-log coordinates. Small dots are profile measurements; large dots are averages for several profiles with similar spreading rates. The dashed line is the best-fit power-law dependence of r.m.s. roughness on spreading rate (roughness = $1,296 \times \text{rate}^{-0.539}$, where roughness is in metres and full spreading rate is in mm per yr).

only a simplified model, because the spatial distribution of volcanic features is bound to be more complex than that of Fig. 4b, but in fact, the same proportionality between h and r.m.s. roughness is obtained using a model that generates a more realistic volcanic layer²³ (Fig. 4c). It should also be pointed out that, although the average height and width of the volcanoes may be quite small (few tens of metres and few hundreds of metres respectively), numerical simulations indicate that the

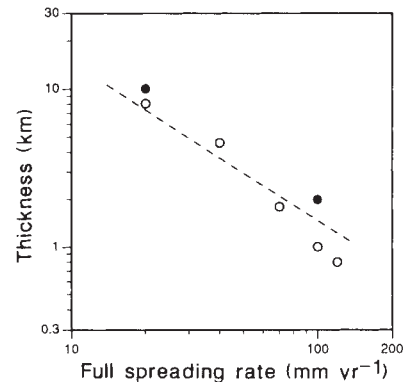


FIG. 3 Estimates of lithospheric thickness at mid-ocean-ridge axes given by the depth to an isotherm (●), depth to the 600 °C isotherm²⁰, ○, depth to the 750 °C isotherm²² are proportional to the inverse of the spreading rate, as indicated by the dashed line.

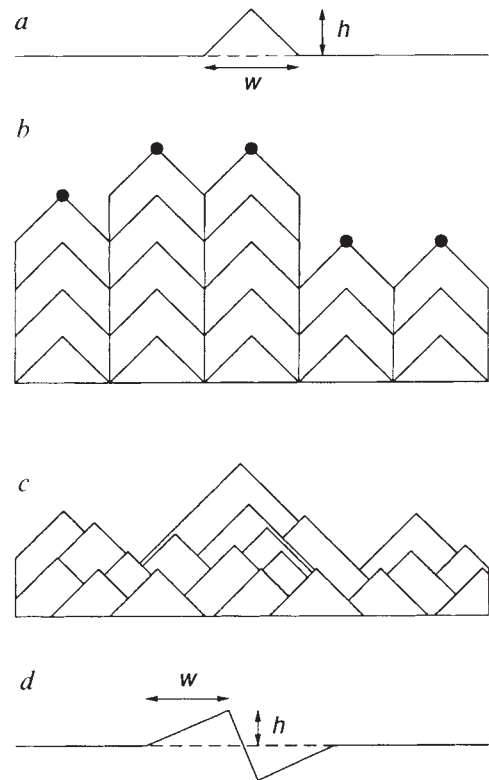


FIG. 4 *a*, Topographic roughness of abyssal hills may be due to the superposition of a number of volcanoes of height h and width w . *b*, In a simple case, the volcanoes are stacked at random but discrete locations; at each of these discrete locations (marked by black dots), there will be a Poisson distribution of the heights, with a r.m.s. roughness proportional to the square root of h . *c*, In a more realistic case, the location of the volcanoes is a continuous random variable, and the surface slopes are limited by the volcano slopes²³; numerical experiments show that the r.m.s. roughness is still proportional to the square root of h , as in the simple case (*b*). *d*, Topographic roughness may also be due to the combined effects of normal faults, where the response of topography to a fault has a half-width w and a relief h .

superposition of a number of these constructions results in large-scale abyssal hills (hundreds of metres high and several kilometres wide)^{23–25}.

Volcano height can be related to the local lithospheric thickness as suggested by Vogt²⁶. If lithospheric thickness and depth of the magma chamber are determined by the thermal structure, then as magma should rise isostatically to a height proportional to the lithostatic pressure at the depth of the magma chamber, volcano height should be linearly proportional to lithospheric thickness. Therefore, the topographic roughness due to the superposition of volcanic constructions is expected to be proportional to the square-root of lithospheric thickness, as predicted.

Explaining the connection between the roughness due to faulting and the lithospheric thickness is more difficult. Consider an idealized fault with a half-vertical displacement h and a half-width of the topographic response w (Fig. 4d). Such a fault will contribute an amount of variance (the square of the r.m.s. roughness) proportional to h^2w , and N faults will result in a variance proportional to h^2wN . It is interesting to note that the work done against gravity to generate the topography due to a single fault²⁷ will be also proportional to h^2w ; this work must be performed by some extension force, and must be equal to the force times the extensional strain. Although both h and w for normal faults near mid-ocean-ridge axes are likely to be greater where the lithosphere is thicker, the form of this relationship cannot be quantified without some poorly constrained assumptions, particularly on how the extension force relates to the lithospheric thickness and on how strain is partitioned between faults.

Although these reasonings are based on several important simplifying assumptions, they are consistent with the idea that the lithospheric thickness near mid-ocean-ridge axes plays an important part in determining the topographic roughness of abyssal hills on the ridge flanks. This may be a valid explanation also for areas originally excluded from this study, where the ridge axis is exceedingly shallow or deep. For example, it has been proposed that the axis of the slow-spreading Reykjanes Ridge is anomalously shallow and is marked by a topographic high because of high mantle temperatures²², so that the axial lithospheric thickness is expected to be less than on the nearby Mid-Atlantic Ridge: in agreement with the discussion above, the flanks of the Reykjanes Ridge are indeed smoother than those of the Mid-Atlantic Ridge^{8,28}. □

Parallel within-island microevolution of lizards on neighbouring islands

R. P. Brown^{*†}, R. S. Thorpe^{*} & M. Báez[‡]

^{*} Zoology Department, University of Aberdeen, Aberdeen AB9 2TN, UK

[‡] Departamento de Zoología, Facultad de Biología, Universidad de La Laguna, La Laguna, Tenerife, Canary Islands, Spain

THE correlation of changes in morphological traits with environmental gradients is often taken as evidence for natural selection¹. But in many cases, the possibility that this correlation is coincidental cannot be ruled out². Stronger evidence for natural selection is provided when closely related allopatric species show parallel patterns of geographic variation along similar environmental gradients. Vertebrate geographic variation within small islands provides a valuable opportunity to detect mechanisms of natural selection^{3–5}. We have now studied microgeographic variation in the colour patterns of two skinks (Lacertilia: Scincidae) from neighbouring heterogeneous islands. Parallel north–south variation occurs in the colour patterns of both species. Populations from the south of both islands possess conspicuous dorsal–tail coloration. Morphological distance matrices for both species are compared with similar aspects of environmental variation using simultaneous Mantel tests. This indicates that differential selection between lush and arid habitats is the primary cause of the variation in colour pattern.

Chalcides sexlineatus is endemic to the Canary Island of Gran Canaria (Spain) where it shows within-island geographic variation in several morphological characteristics^{6,7}. The sister species *Chalcides viridanus* is present on the adjacent island of Tenerife. The colour patterns of 432 *C. viridanus* from 17 localities within Tenerife and 691 *C. sexlineatus* from 47 localities in Gran Canaria were recorded using macrophotography.

Four unit characters quantified the main features of the colour-pattern variation in *C. viridanus*: (1) number of dorsal ocelli; (2) blue tail coloration; (3) green tail coloration; (4) dorsal hue. The first two of these characters could take several different values. Two-way analyses of variance (ANOVA) showed that they varied significantly between localities ($F = 4.48$, $P < 0.001$; $F = 35.81$, $P < 0.001$, respectively) and between adult and juvenile age classes ($F = 4.48$, $P < 0.02$; $F = 54.98$, $P < 0.001$, respectively); a logarithmic transformation of the character value, did not affect these probabilities). The other two characters are either present or not, and their values were not the same between the localities for adults (Pearson χ^2 are 130.46 and 82.76, $P < 0.001$, respectively). The values of these characters were invariant for juveniles. A principal components analysis (PCA) was computed on the normalized group-means of the values of the four adult character to give the generalized geographic variation in colour pattern⁸. The variation is predominantly from north to south (Fig. 1a).

Eight different continuous characters were used to quantify the main features of the colour-pattern variation in *C. sexlineatus*: (1) blue belly pigment; (2) blue dorsal–tail coloration; (3) orange belly pigment; (4) dorsal stripes; (5) number of dorsal ocelli; (6) size of light shoulder blotch; (7) size of lateral dark band at shoulder; (8) ventral black pigment. Two-way ANOVAS showed that the values of the characters vary significantly among localities (F values, respectively: 87.32; 90.76; 37.59; 82.45; 70.97; 35.55; 63.77; 19.56; $P < 0.001$ for all), but are not sexually dimorphic ($P > 0.05$ for all). A logarithmic transformation of the character value did not affect these probabilities. Canonical

Received 7 January; accepted 20 May 1991.

- Menard, H. W. *Science* **157**, 923–924 (1967).
- Lonsdale, P. *Mar. geophys. Res.* **3**, 251–293 (1977).
- Van Andel, T. H. & Bowin, C. O. *J. geophys. Res.* **73**, 1279–1298 (1968).
- Macdonald, K. C. *Ann. Rev. Earth planet. Sci.* **10**, 155–190 (1982).
- Menard, H. W. & Mammertick, J. *Earth planet. Sci. Lett.* **2**, 465–472 (1967).
- Macdonald, K. C. & Atwater, T. M. *Earth planet. Sci. Lett.* **39**, 319–327 (1978).
- Pockalny, R. A., Detrick, R. S. & Fox, P. J. *J. geophys. Res.* **93**, 3179–3193 (1988).
- Talwani, M., Windisch, C. C. & Langseth, M. G. *J. geophys. Res.* **76**, 473–517 (1971).
- Weissel, J. K. & Hayes, D. E. *J. geophys. Res.* **79**, 2579–2587 (1974).
- Bell, T. H. *Deep-Sea Res.* **A26**, 65–76 (1979).
- Fox, C. G. & Hayes, D. E. *Rev. Geophys. Space Phys.* **23**, 1–48 (1985).
- Gilbert, L. E. & Malinverno, A. *Geophys. Res. Lett.* **15**, 1401–1404 (1988).
- Goff, J. A. & Jordan, T. H. *J. geophys. Res.* **23**, 13589–13608 (1988).
- Shaw, P. R. & Smith, D. K. *J. geophys. Res.* **95**, 8705–8733 (1990).
- DeMets, C., Gordon, R. G., Argus, D. F. & Stein, S. *Geophys. J. Int.* **101**, 425–478 (1990).
- Small, C. & Sandwell, D. T. *J. geophys. Res.* **94**, 17383–17392 (1989).
- Small, C., Sandwell, D., Chen, Y. & Royer, J.-Y. *Eos Trans. Am. geophys. Un.* **70**, 1301 (1989).
- Tappinier, P. & Francheteau, J. *J. geophys. Res.* **83**, 3955–3970 (1978).
- Cochran, J. R. *J. geophys. Res.* **84**, 4713–4729 (1979).
- Phipps Morgan, J., Parmentier, E. M. & Lin, J. J. *J. geophys. Res.* **92**, 12823–12836 (1987).
- Solomon, S. C., Huang, P. Y. & Meinke, L. *Nature* **334**, 58–60 (1988).
- Chen, Y. & Morgan, W. J. *J. geophys. Res.* **95**, 17583–17604 (1990).
- Malinverno, A. *Eos Trans. Am. geophys. Un.* **70**, 1302 (1989).
- Malinverno, A. *Eos Trans. Am. geophys. Un.* **72**, 262 (1991).
- Malinverno, A. *Geol. Soc. Am. Mem.* (in the press).
- Vogt, P. R. *Earth planet. Sci. Lett.* **23**, 337–348 (1974).
- Forsyth, D. W. *Eos Trans. Am. geophys. Un.* **71**, 631 (1990).
- Searle, R. C. & Laughton, A. S. *Proc. 26th Int. geol. Congress Oceanologica Acta suppl.* **4**, 5–13 (1981).

ACKNOWLEDGEMENTS. I thank R. Buck, W. Haxby, D. Hayes and N. Mitchell for discussions and reviews of the manuscript and D. Sandwell for providing digitized plate boundaries. This research was supported by the Office of Naval Research.

† Present address: Departamento de Zoología, Facultad de Biología, Universidad de Salamanca, 37071 Salamanca, Spain.

TABLE 1 Variable coefficients for the first canonical variates (*C. sexlineatus*)

	Colour-pattern characters							
	1	2	3	4	5	6	7	8
Males	-0.59	-0.55	0.45	-0.67	-0.07	0.26	0.42	-0.39
Females	-0.71	-0.68	0.52	-0.72	-0.08	0.24	0.41	-0.29

All characters show substantial loadings on the first canonical variates (except character 5). For *C. viridanus*, normalized eigenvector coefficients for the first principal component were 0.55, -0.62, 0.01 and 0.56 for characters 1, 2, 3 and 4, respectively. The rationale for computing canonical variates and principal components analyses was that they condense most of the among-locality variation into (in this case) a single vector which can then be conveniently portrayed using a contouring program (Fig. 1).

variate analyses (CVA) computed for the eight characters condensed most of the between-group variation in *C. sexlineatus* (males, 72.3%; females 77.8%) into one measure of colour-pattern variation, that is, the first canonical variate (Table 1) (the results were checked using PCA to ensure that heteroscedasticity had not perturbed the CVA^{3,9}). This variable shows a clear NNE-SSW pattern of geographic variation (Fig. 1b).

Three hypotheses could explain the colour-pattern variation in *C. sexlineatus* on Gran Canaria, whereas two hypotheses are feasible for *C. viridanus* on Tenerife. These are:

■ **Lush/arid habitats hypothesis** (*C. viridanus* and *C. sexlineatus*, Fig. 2). There are pronounced differences in climate and vegetation between the north and south of both Tenerife and Gran Canaria. These seem to be the primary cause of geographic variation in other lizard species on the former island^{3,4,5}. The categorical nature of the lush-arid variation allows the use of a dual-state variable for hypothesis testing.

■ **Altitude hypothesis** (*C. sexlineatus* only; *C. viridanus* is a predominantly low-altitude species). Gran Canaria reaches a height of 1,949 m, giving rise to altitude-related variation in climate and vegetation^{10,11}. The scalation of several species of lizards has been shown to covary with altitude^{4,6}. Localities were therefore scored according to their height above sea level.

■ **Geographical proximity hypothesis** (*C. viridanus* and *C. sexlineatus*), which models gene flow according to an isolation-by-distance model¹².

A two-hypothesis simultaneous Mantel test was used to test the models¹³. The colour-pattern distance matrix for each species consisted of the mean population scores for all colour characters

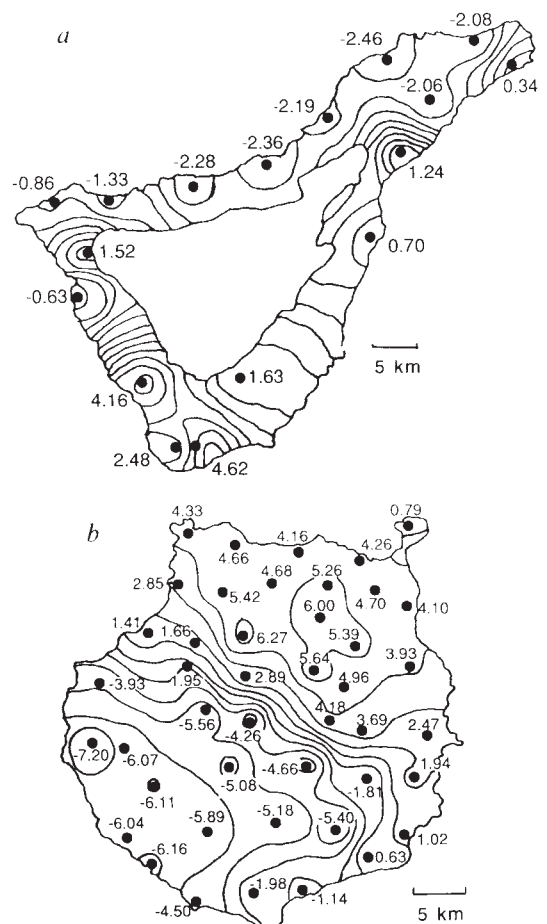


FIG. 1 Patterns of geographic variation in coloration. a, Contour plot of the first principal component representing 58.3% of total among-population variation in *C. viridanus* in Tenerife (the blank centre is the area of 'absence'). b, The same for female *C. sexlineatus* on Gran Canaria (first canonical variate representing 77.8% of total among-population variation).

('observed' matrices). The two 'hypothesized' matrices represented the lush/arid habitats hypothesis or altitude hypothesis in addition to the geographical-proximity hypothesis.

The lush/arid habitats hypothesis is significant for both adult and juvenile *C. viridanus* (Table 2). For male and female *C. sexlineatus*, the lush/arid habitats hypothesis is significant and

TABLE 2 Mantel tests

Colour-pattern matrix (1)	Hypothesized matrices		$g_{1,2}$	$g_{1,3}$	$g_{1,2-3}$	R^2
	(2)	(3)				
<i>C. viridanus</i> (juveniles)	Lush/arid habitats	Geographical proximity	0.142 **	0.269 ***	0.117 *	0.142 **
<i>C. viridanus</i> (adults)	Lush/arid habitats	Geographical proximity	0.164 **	0.259 ***	0.141 **	0.219 ***
<i>C. sexlineatus</i> (males)	Lush/arid habitats	Geographical proximity	0.368 ***	0.276 ***	0.307 ***	0.411 ***
<i>C. sexlineatus</i> (females)	Lush/arid habitats	Geographical proximity	0.360 ***	0.274 ***	0.296 ***	0.400 ***
<i>C. sexlineatus</i> (males)	Altitude	Geographical proximity	0.011 NS	0.274 ***	0.056 NS	0.243 ***
<i>C. sexlineatus</i> (females)	Altitude	Geographical proximity	0.004 NS	0.277 ***	0.052 NS	0.261 ***

The five statistics computed by the Mantel tests are (1) a simple regression ($g_{1,2}$) between the second hypothesized and the observed matrices, (2) a simple regression ($g_{1,3}$) between the second hypothesized and the observed matrices; (3) a partial regression ($g_{1,2-3}$), of the observed against the first hypothesized matrices with the effects of the second hypothesized matrix (proximity) removed; (4) a squared multiple correlation (R^2). The values for these statistics are then compared with the ranges of values for the same statistics given by 500 random associations of the matrices. * $P < 0.1$; ** $P < 0.01$; *** $P < 0.002$; NS (not significant), $P > 0.1$.

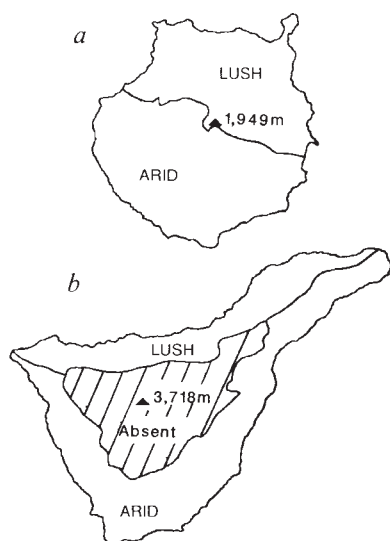


FIG. 2 Lush/arid habitat/hypothesis modelling habitat variation within the islands. The island's topography and the prevailing north/north-east trade winds, result in cloud formation on the north-facing slopes^{10,11}. By contrast, the south of the islands are generally cloud-free, exhibiting higher temperatures and lower rainfall. These latitudinal climatic transitions are pronounced and give rise to corresponding changes in vegetation. For example, the original extent of the luxuriant laurisilva forest coincides with areas of semi-permanent cloud cover on the north-facing slopes^{16,11}. Vegetation cover in the south is extremely sparse, consisting of xerophytic plants. Well-defined ecotones separate 'lush' and 'arid' biotopes, the exact location of which were determined by aerial photographs of the islands (see, for example, ref. 17). *a*, Gran Canaria. *b*, Tenerife. (Note that *C. viridanus* seems to be absent or only present at extremely low population densities above 1,100 m on Tenerife, so this part of the island (hatched) was not included when calculating the hypotheses distance matrices.)

remains so after the removal of geographical proximity. The altitude hypothesis is rejected.

These findings indicate that the lush/arid changes in climate and vegetation within the islands may determine natural selection for different colour-pattern morphs. Possibly the most interesting aspect of the variation is the presence of a conspicuous blue-green dorsal-tail coloration in animals from the arid southern areas of both of the islands. This could be an adaptive response to predation pressure from kestrels (*Falco tinnunculus*) and perhaps shrikes (*Lanius excubitor*). Tail autotomy, the ability to shed the tail when attacked, is a well-known anti-predator device in lizards¹⁴, the effectiveness of which may be increased by conspicuous tail coloration¹⁵. Skinks from the north of both islands have generally dark, uniform, cryptic dorsal colour patterns. It seems likely that these morphs represent alternative anti-predator strategies which have arisen by the different selective pressures that exist between lush and arid habitats. Different predator escape strategies could be favoured in different habitats¹⁴ (that is crypsis versus conspicuous tail strategies). Alternatively, the shrike is thought to be more abundant in arid habitats, which could reflect geographic variation in predation pressure. □

Received 28 January; accepted 13 May 1991.

1. Endler, J. A. *Geographic Variation, Speciation and Clines* (Princeton University Press, Princeton, New Jersey, 1977).
2. Endler, J. A. *Natural Selection in the Wild* (Princeton University Press, Princeton, New Jersey, 1986).
3. Thorpe, R. S. *Syst. Zool.* (in the press).
4. Thorpe, R. S. & Báez, M. *Evolution* **41**, 256–268 (1987).
5. Thorpe, R. S. & Brown, R. P. *Biol. J. Linn. Soc.* **38**, 303–322 (1989).
6. Brown, R. P. & Thorpe, R. S. *Biol. J. Linn. Soc.* (in the press).
7. Brown, R. P. & Thorpe, R. S. *J. Evol. Biol.* (in the press).
8. Thorpe, R. S. *Biol. J. Linn. Soc.* **14**, 215–233 (1980).
9. Thorpe, R. S. in *Numerical Taxonomy* (ed. Felstein, J.) 404–423 (NATO Advanced Study Institute Series G (Ecological Sciences), No. 1, Springer-Verlag, Berlin, 1983).
10. Fernandopulle, D. in *Biogeography and Ecology of the Canary Islands* (ed. Kunkel, G.) 185–206 (Junk, The Hague, The Netherlands, 1976).

11. García, J. L. Hernández, J. Cabrera, L. G. Díaz, A. & Afonso L. *Atlas Interinsular de Canarias* (Editorial Interinsular Canaria, S.A. Tenerife, Spain, 1990).
12. Douglas, M. E. & Endler, J. A. *J. theor. Biol.* **99**, 777–795 (1982).
13. Manly, B. F. J. *Res. Pop. Ecol.* **28**, 201–218 (1986).
14. Arnold, E. N. *J. Nat. Hist.* **18**, 127–169 (1984).
15. Cooper, W. E. & Vitt, L. J. *Z. Tierpsychol.* **70**, 265–276 (1986).
16. Kammer, F. *Scripta Geobotanica* **7**, 1–78 (1974).
17. C.N.E.S. & Maptec Int. Ltd. *Imagen Satellite: Gran Canaria* (CNES Quaestus Int. Ltd. (dist.), Dublin, Ireland, 1987).

ACKNOWLEDGEMENTS. This research was funded by the Natural Environment Research Council (R.P.B.), Carnegie Trust (R.S.T.) and the British Council (R.S.T. and M.B.).

Dorsal nostrils and hydrodynamically driven underwater olfaction in plesiosaurs

Arthur R. I. Cruickshank, Philip G. Small* & Michael A. Taylor

Leicestershire Museums, Arts and Records Service, 96 New Walk, Leicester LE1 6TD, UK

* X-Ray Department, Medical School, Queen's Medical Centre, University of Nottingham, Nottingham NG7 2UH, UK

THE dorsally placed external nostrils of plesiosaurs are usually regarded as an adaptation to breathing in those extinct marine reptiles. We suggest instead that the narial system was used in underwater olfaction. The internal nares are anterior to the external nares. Hydrodynamic pressure during swimming forced water into the mouth, along palatal grooves into the scoop-shaped internal nares and up short ducts, presumably lined with olfactory epithelia. Alternatively, or additionally, the so far unlocated Jacobson's organ detected particulate matter. The water was sucked out through the external nares by hydrodynamic pressures generated by fast flow over the convex upper surface of the head.

We used an almost uncrushed skull of the pliosauroid *Rhomaleosaurus megacephalus* (Stutchbury 1846) (Leicestershire Museums, Arts and Records Service G221.1851, Hettangian, Lower Jurassic, Barrow upon Soar, Leicestershire, England^{1,2}). The matrix is clayey limestone with little or no radiodense pyrite. Direct X-radiography and computed axial tomography (Figs 1, 4 and 5) gave good results. The specimen was found to be broken into four major pieces when it was dismounted, and we studied the section containing the internal and external nares (Figs 4 and 5).

The palate is typical of plesiosaurs (Fig. 2). Each premaxilla bears five widely spaced, semi-recumbent teeth with a median gap between the anteriormost teeth. The dentary teeth intermesh with, but do not contact, the upper teeth. Like modern reptiles, plesiosaurs presumably lacked fleshy lips, and the gum material surrounding the roots of the teeth did not allow tight sealing of the jaws and water could therefore flow through the gaps between the teeth.

There is no secondary palate, so little direct comparison with marine mammals is possible^{3,4}. On each side of the palate, two grooves lead backwards from the premaxillary-maxillary diastema (Figs 1 and 2). The inner groove runs to the internal naris, which is inset into a swelling terminating the inner groove. The left internal naris is about 2.5 cm long by 1.0 cm wide. The internal naris strongly resembles the NACA Duct (National Advisory Committee for Aeronautics) design of air scoops on aircraft, whose opening is flush with the surface. This design deepens and widens in the direction of fluid flow and ends in a transverse scoop (Fig. 3), suggesting that the internal naris acted as a ram scoop for water flowing backwards along the inner groove. The outer groove runs to a surprisingly large foramen posterior to the internal naris, which probably housed

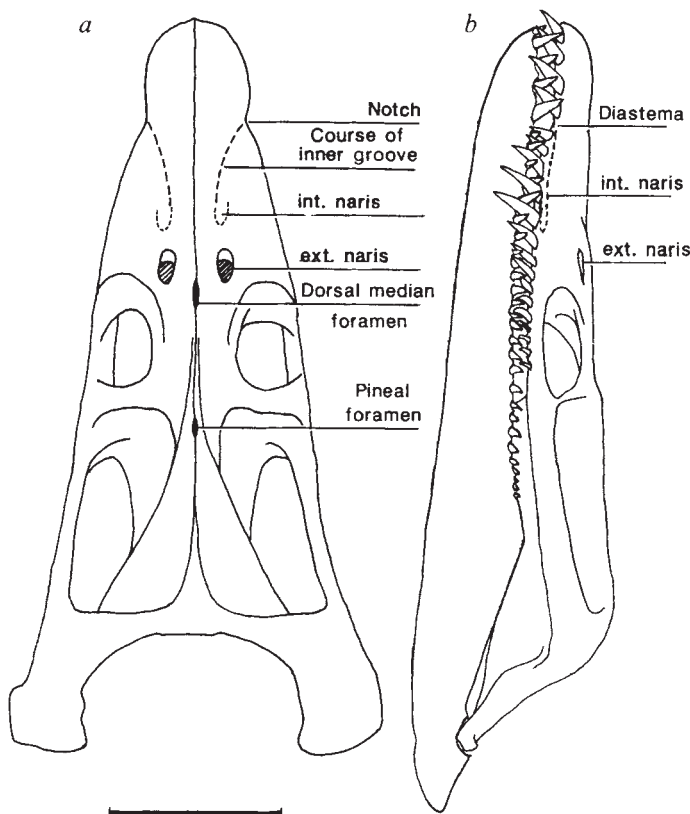


FIG. 1 Reconstruction of upper surface (a) and side view (b) of *Rhomaleosaurus* (M.A.T., manuscript (in preparation)) showing relative position of internal and external nares and 'pinched in' region of the snout. Scale bar, 20 cm. Abbreviations used in the figures: ext. naris, external naris; f, frontal; int. naris, internal naris; l, jaw, lower jaw; max, maxilla; pal, palatine; pmx, premaxilla; pt, pterygoid; vo, vomer.

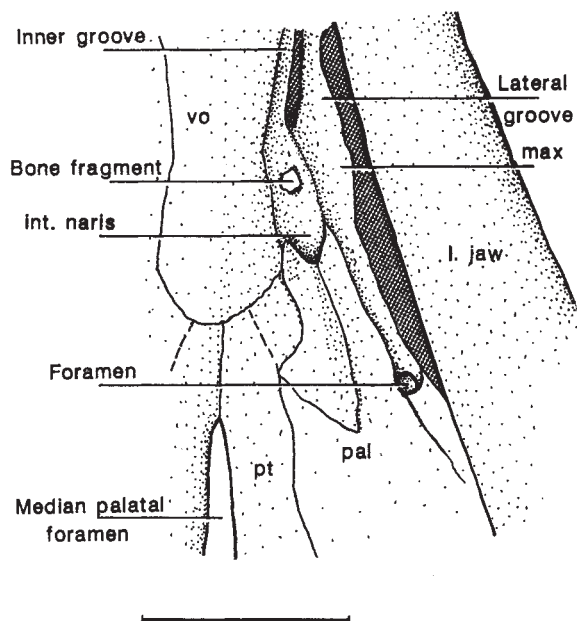


FIG. 2 Detail of the palate of *Rhomaleosaurus megacephalus* (LEICM G221.1851) showing the structure in the region of the internal nares. Stipple, rock matrix. Scale bar, 5 cm.

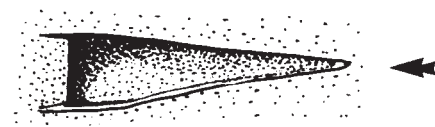


FIG. 3 NACA-type of flush duct. Arrow marks direction of flow. Not to scale.

blood vessels, nerves, or both, supplying the anterolateral palate.

The posterior rim of the naris is a steep wall, above which the duct disappears (Figs 2 and 5). A circular duct, about 1.0 cm in diameter, runs upward from the internal to the front of the external naris, which lies about 4 cm behind the internal naris (Figs 1, 2 and 5). Each external naris is oval, about 4.0 cm long by 1.75 cm wide. The anterior rim of the external naris is steep and the gently sloping posterior rim opens into a broadening groove running posteriorly between the orbit and the median ridge. The external morphology of the external naris indicates that it was an outlet for a posteriorly directed current. We suggest that the fast flow of water over the convex dorsal surface of the head generated negative pressures which sucked water out of the external nares.^{5,6}

The internal nares are much smaller than the external nares and must have had a markedly greater flow velocity. The water would have been slowed before passage over the olfactory organs, as design considerations suggest⁵⁻⁷.

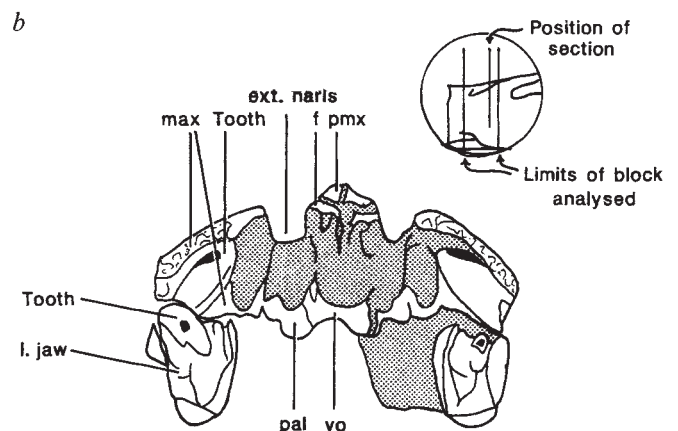
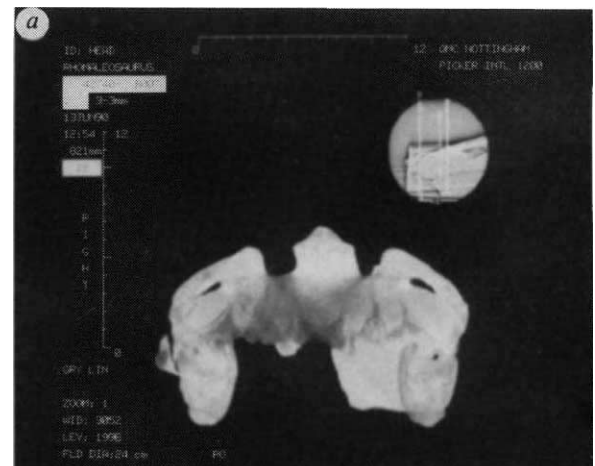


FIG. 4. *a*, Computed Axial Tomography (CAT) scan-section of the narial region of *Rhomaleosaurus megacephalus* (LEICM G221.1851). Photographic print taken from original on 'Picker' Synergview 1200SX, 130 kV and 95 mA. Inset to show the level of section relative to snout region. *b*, Tracing of *a* to show structures seen in the section. Stipple, rock matrix. Scale bars on *a*.

Were the nares used for respiration? Both internal and external nares would have been partly occluded by soft tissue and any measurement of the bony outlines will only give a maximum dimension. In the very similar *R. zetlandicus* (M.A.T., manuscript in preparation), sculpture on the anterior external narial wall may indicate insertions for muscles and ligaments controlling a narial valve. In *R. megacephalus* each internal naris had a functionally clear area of 2.0 cm² or less, which seems disproportionately small for a 5-m-long *Rhomaleosaurus*. Also, the placement of the internal nares anterior to the external nares would have hindered respiratory air flow. We believe that *Rhomaleosaurus* breathed through its open mouth, contrary to the conventional view that all respiration occurs through the narial system. Water would have been excluded from the trachea during normal swimming and in heavy seas, by a glottal valve and a large muscular tongue. The reconstruction of anteriorly hinged nasal flaps on the external nares would presume automatic closure if inspiration were attempted through them. We therefore believe that the dorsal placement of the external nares is thus irrelevant for respiration, unlike in crocodilians and cetaceans, which have full secondary palates. In these two groups with secondary palates, respiration is associated with airborne olfaction, but in the case of this plesiosaur we believe the adaptation to be for the sampling of water and not air^{8,9}.

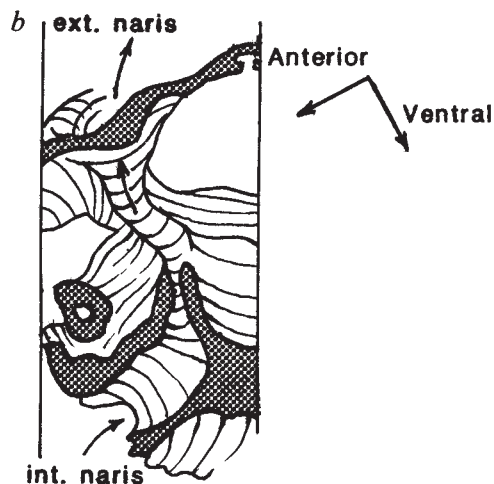


FIG. 5 *a*, Synthetic (three-dimensional) parasagittal section of left nasal region of *R. megacephalus* showing short, straight nasal duct with possible diverticulum. Based on CAT Scanning of specimen LEICM G221.1851. Reconstruction using '3-D Reconstruction Program' of Picker Synview 1200SX. *b*, Tracing of *a* to show structures seen in the three-dimensional reconstruction. Not to scale. Note arrows for orientation. Stipple, cut edge of bone.

Detailed hydrodynamic studies are needed to confirm these suggestions for *Rhomaleosaurus*. Most plesiosaurs had the same relative position of external and internal nares¹⁰ and we speculate that underwater olfaction was primitive, used to detect prey and carrion and for intraspecific communication as in crocodilians¹¹. It is not currently possible to make a convincing cladistic analysis of these and other characters of the plesiosaurs, but further work may provide a more thorough understanding of their origins and relationships.

Broadly similar mechanisms of water flow occur in carcharhinid sharks (though with two pairs of external nares and no internal nares)¹² and should be considered for ichthyosaurs which have dorsal nares but no secondary palate and may also have used underwater olfaction¹³. Finally, such a mechanism allows measurement of velocity through the water by the pressure differentials and water flow through the narial cavities, analogous to the pitot-static system of aircraft, whereas differential indications between left and right nares would indicate yaw and asymmetric flow. □

Received 1 February; accepted 21 May 1991.

1. Taylor, M. A. & Cruickshank, A. R. I. *Trans. Leicestershire lit. phil. Soc.* **83**, 20–24 (1989).
2. Martin, J. G., Frey, E. & Riess, J. *Trans. Leicestershire lit. phil. Soc.* **80**, 58–72 (1986).
3. Lowell, W. R., & Flannigan, W. F. Jr *Mammal Rev.* **10**, 53–59 (1980).
4. Cave, A. J. E. *J. Zool., Lond.* **214**, 307–311 (1988).
5. Vogel, S. *Life in Moving Fluids* (Princeton University Press, Princeton, 1981).
6. Vogel, S. *Am. Scientist* **76**, 28–34 (1988).
7. LaBarbera, M. *Science* **249**, 992–1000 (1990).
8. Atema, J. *Soc. exp. Biol. Symp.* **39**, 387–423 (1985).
9. Atema, J. in *Comparative Physiology: Life in Water and on Land* (eds Dejours, P., Bolis, L., Taylor, C. R. & Weibel, E. R.) 303–316 (Fidia Research Series, IX-Liviana Press, Padova, 1987).
10. Brown, D. S. *Bull. Br. Mus. nat. Hist. Geol.* **35**, 253–347 (1981).
11. Lang, J. W. in *Crocodiles and Alligators* (ed. Ross, C. A.) 102–117 (Merehurst, London, 1989).
12. Zeiske, E., Theisen, B. & Gruber, S. H. *Can. J. Zool.* **65**, 2406–2412 (1987).
13. McGowan, C. *Bull. Br. Mus. nat. Hist. Geol.* **24**, 1–109 (1981).

ACKNOWLEDGEMENTS. We thank the Director of the Leicestershire Museums, Arts and Records Service, J. Martin and C. Collins for support and access to the material, S. Thursfield for the photographs, and J. Noble and J. Bankin for their enthusiastic assistance. A.R.I.C. acknowledges financial aid from the Open University and M.A.T. the support of the University of Leicester.

Extracellular domains mediating ϵ subunit interactions of muscle acetylcholine receptor

Xiao-Mei Yu & Zach W. Hall*

Department of Physiology, School of Medicine, University of California at San Francisco, San Francisco, California 94143-0444, USA

LIGAND-gated ion channels, a major class of cell-surface proteins, have a pseudosymmetric structure with five highly homologous subunits arranged around a central ion pore¹. The correct assembly of each channel, whose subunit composition varies with cell type and stage of development, requires specific recognition between the subunits^{2–4}. Assembly of the pentameric form of the acetylcholine receptor from adult muscle (AChR; $\alpha_2\beta\epsilon\delta$) proceeds by a stepwise pathway starting with the formation of the heterodimers, $\alpha\epsilon$ and $\alpha\delta$. The heterodimers then associate with the β subunit and with each other to form the complete receptor^{5–7,21}. We have now determined which parts of the subunits mediate the interactions during assembly of the adult form of the receptor from mouse muscle by using a chimaeric subunit in which the N-terminal and C-terminal extracellular domains are derived from the ϵ subunit with the remainder from the β subunit. The ϵ and β subunits were chosen because the ϵ subunit forms a heterodimer with the α subunit in the pathway for assembly of the receptor,

* To whom correspondence should be addressed.

whereas the β subunit does not. The ϵ_β chimaera can substitute for the ϵ but not the β subunit in the oligomeric receptor, indicating that the α subunit specifically recognizes an extracellular domain of the ϵ subunit.

To make chimaeras with ϵ and β , new restriction sites were introduced into both subunit complementary DNAs at the 5' end of M1 (*Sna*BI) and the 3' end of M4 (*Spe*I) by oligonucleotide-directed mutagenesis (Fig. 1). These sites were chosen to separate the major extracellular domains at the N and C termini of each subunit from the intervening transmembrane, intracellular and extracellular domains. The segments between the new restriction sites were then exchanged between the two subunit cDNAs by conventional methods to generate two complementary chimaeric subunit cDNAs: ϵ_β , whose termini are derived from the ϵ subunit; and β_ϵ , with termini from the β subunit (Fig. 1).

COS cells transfected with a combination of cDNAs for α , β , ϵ and δ subunits expressed the acetylcholine receptor on their surface (Table 1); this receptor has properties that are similar to those of the acetylcholine receptors at adult endplates⁸. When either the β or ϵ subunit cDNA was omitted from the transfection mixture, no toxin-binding activity was observed on the cell surface, indicating that acetylcholine receptors were not expressed there (Table 1). We then tested the ability of the chimaeric cDNAs to restore surface toxin-binding activity to cells transfected with the deficient mixtures. We found that the cDNA encoding ϵ_β could substitute for the ϵ , but not the β , subunit cDNA in the transfection assay. Addition of ϵ_β cDNA to a transfection mixture containing α , β and δ cDNAs increased the toxin-binding activity to roughly 30% of that obtained with the complete mixture, whereas ϵ_β added to a mixture containing α , ϵ and δ cDNAs had no effect (Table 1). The β_ϵ subunit cDNA failed to produce surface toxin-binding activity in either case. Chimaeras in which only the ϵ N- or C-terminal domain was substituted into the β subunit also failed to yield surface toxin-binding activity. Because chimaeric proteins often fold incorrectly^{9,10}, we could not make conclusions about the specificity of subunit interactions for chimaeric subunits that did not give rise to surface toxin-binding activity; we did not investigate these further.

The relative effectiveness of ϵ_β and ϵ subunits in promoting expression of toxin-binding activity was compared by adding increasing amounts of each to transfection mixtures containing a fixed amount of α , β and δ cDNAs. Complementary DNAs for the ϵ_β and ϵ subunits were equally effective at low cDNA concentrations, but at higher concentrations the amount of surface toxin-binding activity with ϵ cDNA was about twice that with ϵ_β cDNA (Fig. 2a).

The toxin-binding complex apparently contains all four subunits. Thus, the expression of toxin-binding activity on the

surface of COS cells transfected with α , β , δ and ϵ_β subunits was reduced by over 85% by the omission of either ϵ_β , β or δ subunit cDNAs. In addition, the surface toxin-binding complex was immunoprecipitated by subunit-specific antibodies to α , β and δ subunits (data not shown). The antibody to the β subunit (monoclonal antibody 124), which recognizes an epitope on the cytoplasmic loop connecting M3 and M4 (ref. 11), stained both β and ϵ_β subunits in immunoblots of total cell extracts (data not shown), and so recognizes both. Consistent with its being a pentamer, sucrose gradient sedimentation showed that the complex has a sedimentation constant of 9.5S, as does the native acetylcholine receptor (Table 2).

Many of the other properties of the toxin-binding complex formed by ϵ_β are also similar to those of the acetylcholine receptor. The association rate constant of the toxin-binding

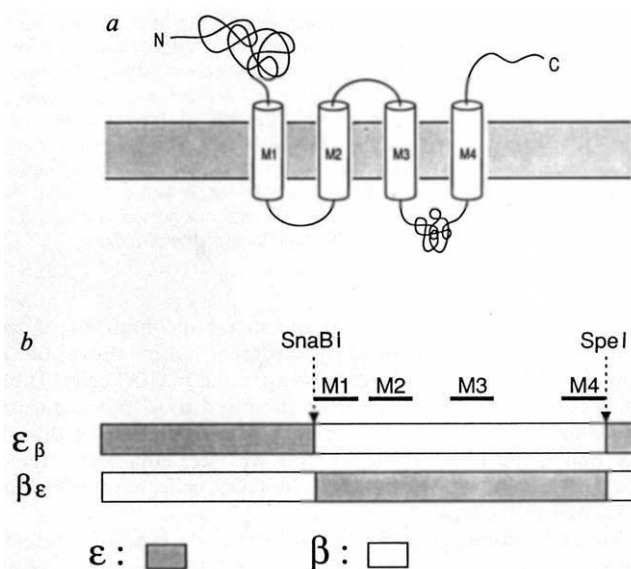


FIG. 1 a, A general model of the transmembrane topology of acetylcholine receptor subunits proposed on the basis of primary sequence hydrophobicity^{16,17}. b, Structure of chimaeric ϵ_β and β_ϵ subunits. The structures of chimaeric ϵ_β and β_ϵ subunits are diagrammatically shown with the hydrophobic segments M1-M4 aligned, and with the length difference between them (8 amino acids) ignored. Open boxes, mouse β -subunit amino-acid sequences; filled boxes, mouse ϵ -subunit amino-acid sequences¹⁸. The locations of the introduced *Sna*BI and *Spe*I sites are indicated. The chimaeric ϵ_β subunit is made up of ϵ residues -20-222, β residues 225-461 and ϵ residues 452-473, whereas the chimaeric β_ϵ subunit is made up of β residues 23-224, ϵ residues 223-451 and β residues 462-478.

METHODS. *Sna*BI and *Spe*I restriction sites were introduced into homologous locations in pSM β and pSM ϵ by oligonucleotide-directed mutagenesis^{19,20}. Uracil-containing single-strand DNAs of pSM β and pSM ϵ were rescued from CJ₂₃₆ by using M₁₃ helper phage. The responsible oligonucleotides overlapped the wild-type sequence for 10 bases on either side of the last base(s) changed. In pSM β , the nucleotide sequence TACCTG (encoding the β residues Tyr 225 and Leu 226) was mutagenized to TACGTA (Tyr 225 and Val 226) to introduce the *Sna*BI restriction site. The changes in the pSM β nucleotide sequence required for the introduction of the *Spe*I site did not alter the primary β sequence of Thr 461, Leu 462 and Val 463 (original, ACCCTAGTC; mutant, AACTAGTC). In pSM ϵ , the *Sna*BI site was introduced into the nucleotide sequence coding for the ϵ residues Tyr 222, Val 223 by changing TACGTC to TACGTA without primary sequence alterations, whereas the *Spe*I site was created by changing the nucleotide sequence TCTACTCTC (corresponding to ϵ residues Ser 450, Thr 451, Leu 452) into TCACTAGTC (ϵ residues Ser 450, Leu 451, Val 452). When either the pSM β or pSM ϵ construct containing the *Sna*BI and *Spe*I sites was cotransfected into COS cells with the other acetylcholine receptor subunit cDNAs, wild-type levels of ¹²⁵I-labelled BTX binding to the surface was detected (data not shown). After the mutated pSM β and the pSM ϵ cDNAs were digested with *Sna*BI and *Spe*I, the resulting 0.7-kilobase (kb) fragments, encoding the region between the start of M1 and the end of M4, were purified and ligated into the purified vectors from the other construct, containing the N and C termini, to generate the desired chimaeric subunits.

TABLE 1 ¹²⁵I-labelled BTX binding on the surface of COS cells

Subunits expressed	Surface acetylcholine receptor (fmol per well)
$\alpha\beta\epsilon\delta$	46.7 ± 13.3
$\alpha\beta\delta$	1.0 ± 0.6
$\alpha\beta(\epsilon_\beta)\delta$	13.6 ± 0.6
$\alpha\beta(\beta_\epsilon)\delta$	0.3 ± 0.4
$\alpha\epsilon\delta$	0.8 ± 1.1
$\alpha(\epsilon_\beta)\epsilon\delta$	0.6 ± 0.1
$\alpha(\beta_\epsilon)\epsilon\delta$	0.2 ± 0.8

COS cells in 60-mm dishes with 50% confluency were transfected with various combinations of mouse acetylcholine receptor subunit cDNAs as shown. The transfected cells were replated to four wells in a 24-well plate 24 h after the transfection and were incubated with ¹²⁵I-labelled bungarotoxin (α -BTX) 24 h later. ¹²⁵I-labelled (α -BTX) binding to the intact cells was determined as previously described⁸. Each value is the mean ± s.e.m. of three determinations.

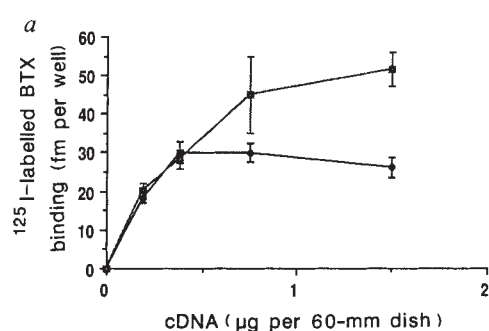
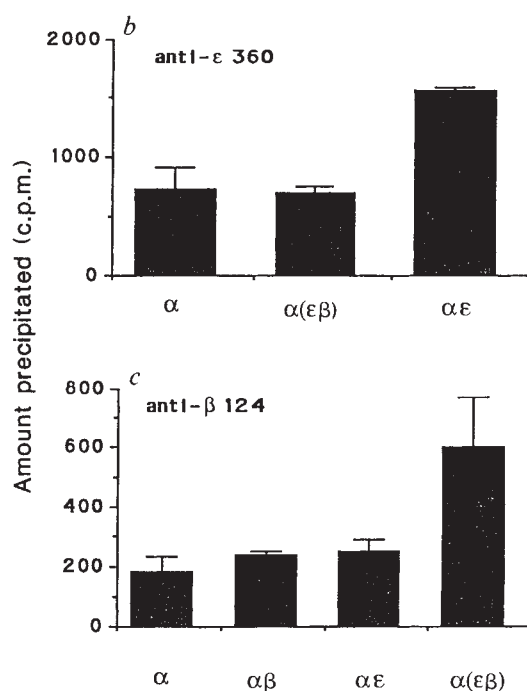


FIG. 2 *a*, Expression of toxin-binding activity as a function of the amount of ϵ (■) or $\epsilon\beta$ (◆) cDNA. COS cells in 60-mm dishes were transfected with fixed amounts of α (1.32 μ g), β (0.66 μ g) and δ (0.26 μ g) cDNAs plus the indicated amount of ϵ or $\epsilon\beta$ subunit cDNA. For each concentration, the cells were then trypsinized and replated into three wells of a 24-well cluster plate, and the surface binding of 125 I-labelled α -BTX determined 24 h later. Each point represents the mean $\pm$ s.e.m. of three determinations. *b*, Association of mouse ϵ subunit and chimaeric $\epsilon\beta$ subunit with α subunit. *c*, Association of ϵ subunit with α subunit. COS cells transfected with subunit cDNAs as indicated were lysed in a 10-cm dish and the cell lysates incubated with 125 I-labelled α -BTX to label toxin-binding sites. Toxin-binding activity was then immunoprecipitated as described in ref. 3 with anti- ϵ 360 antibodies, which are specific for ϵ subunit, or with monoclonal antibody 124, which is specific for the large cytoplasmic loop of the β subunit.



reaction for the complex and the metabolic half-life of the complex are not significantly different from those of ϵ -containing acetylcholine receptor expressed in COS cells (Table 2). Also, the binding of toxin is inhibited by *d*-tubocurarine; the complex is even more sensitive to *d*-tubocurarine than is the native receptor (Table 2). But we were unable to detect acetylcholine-sensitive channels in COS cells expressing the $\epsilon\beta$ -containing complex.

Taken together, the properties of the toxin-binding complex formed when $\epsilon\beta$ is present suggest that the complex has an oligomeric structure analogous to that of the acetylcholine receptor and that it contains the $\epsilon\beta$ subunit instead of the ϵ subunit. The complex is recognized as a complete receptor by the cell, which transports it to the surface, and degrades it with a normal turnover time. In other experiments, we have shown that incompletely or incorrectly assembled subunits are not transported to the surface of COS cells, but are retained intracellularly (ref. 21, and J. R. Forsayeth and Z.W.H., unpublished experiments).

To examine further the role of the chimaeric $\epsilon\beta$ subunit in the pathway of assembly, we investigated whether it forms a heterodimer with the α subunit. COS cells were each transfected with α subunit cDNA and ϵ , β or $\epsilon\beta$ subunit cDNA. Extracts of the transfected cells were then incubated with 125 I-labelled α -bungarotoxin and immunoprecipitated either with affinity-

purified antibodies for the ϵ subunit, or with the monoclonal antibody 124 that recognizes the long cytoplasmic loop of the β subunit. Immunoprecipitation of toxin-binding activity was taken as evidence of heterodimer formation. Both ϵ and $\epsilon\beta$ subunits, but not β , formed a heterodimer with α (Fig. 2*b, c*). Thus, substitution of the N- and C-terminal domains of the β subunit with those of the ϵ subunit enables it to form a specific complex with the α subunit. As the formation of such a complex is the first step in the pathway of the acetylcholine receptor assembly^{6,7,21}, the ability of $\epsilon\beta$ to associate with α in the absence of other subunits presumably underlies its ability to be assembled into an oligomeric receptor and subsequently transported to the surface.

That the terminal extracellular domains of the ϵ subunit determine the specificity of subunit interactions suggests that the subunits of the acetylcholine receptor recognize each other by direct interactions among these domains. Indeed, the relative efficiency of rat and mouse ϵ subunits in promoting association with α subunits in COS cells is determined by two amino acids in the N-terminal extracellular domain⁵. In addition, a chimaeric subunit similar to $\epsilon\beta$, but with terminal domains from the δ subunit, also forms a heterodimer with the α subunit and supports the expression of the acetylcholine receptor, replacing δ in the assembly pathway. This suggests that recognition by one or both terminal domains may be a general mechanism involved in the assembly of receptors from subunits (unpublished data).

The subunits of several transmembrane oligomeric proteins with simple structures, such as the haemagglutinin protein and neuraminidase of influenza virus^{12,13}, the insulin receptor¹⁴ and the ($\text{Na}^+ + \text{K}^+$)ATPase (K. J. Renaud, E. M. Inman and D. M. Fambrough, personal communication), interact through their extracellular domains. By contrast, interactions between two subunits of the T-cell receptor complex are mediated by transmembrane domains¹⁵. In the acetylcholine receptor the hydrophobic domains of the subunits are closely associated to form the ion channel¹. Our experiments clearly show, however, that the extracellular domains carry out a recognition step that is necessary for assembly of this complex oligomeric protein, a step that could occur before association of the hydrophobic domains to form the channel. Further definition of the extracellular domains involved in recognition may lead to understanding of the mechanisms that determine the identity and

TABLE 2 Properties of the ϵ or $\epsilon\beta$ -containing acetylcholine receptor on the surface of COS cells

	ϵ -containing receptor	$\epsilon\beta$ -containing receptor
Sedimentation coefficient (S)*	9.5	9.5
Association rate constant for α -BTX ($10^4 \text{ M}^{-1} \text{ s}^{-1}$)†	2.3 ± 0.1	2.6 ± 0.2
IC ₅₀ for <i>d</i> -tubocurarine (M)‡	2.5×10^{-7}	3.0×10^{-10}
Metabolic half life of receptor (h)‡	23	19

Values were determined in each case as described⁸. IC₅₀, concentration of *d*-tubocurarine at which there is 50% inhibition of α -BTX binding is observed.

* Single determination.

† Mean $\pm$ s.e.m. of six determinations from two separate experiments.

‡ Average of two determinations.

relative position of subunits in different members of the ligand-gated receptor family and in other oligomeric membrane proteins. □

Received 19 February; accepted 15 May 1991.

- Unwin, N. *Neuron* **3**, 665–676 (1989).
- Mishina, M. *et al. Nature* **321**, 406–411 (1986).
- Gu, Y. & Hall, Z. W. *Neuron* **1**, 117–125 (1988).
- Betz, H. *Neuron* **5**, 383–392 (1990).
- Gu, Y. *et al. Neuron* (in the press).
- Blount, P. & Merlie, J. P. *Neuron* **3**, 349–357 (1989).
- Blount, P. *et al. J. Cell Biol.* **111**, 2601–2611 (1990).
- Gu, Y. *et al. Neuron* **5**, 147–157 (1990).
- Rose, J. K. & Doms, R. W. *A. Rev. Cell Biol.* **4**, 257–288 (1988).
- Hurtley, S. M. & Helenius, A. *A. Rev. Cell Biol.* **5**, 277–307 (1989).
- Gullick, W. & Lindstrom, J. *Biochem.* **22**, 3312–3320 (1983).
- Doyle, C. J. *et al. J. Cell Biol.* **103**, 1193–1204 (1986).
- Varghese, J. N. *et al. Nature* **303**, 35–40 (1983).
- Johnson, J. D. *et al. Proc. natn. Acad. Sci. U.S.A.* **85**, 7516–7520 (1988).
- Manolios, N. *et al. Science* **249**, 274–277 (1990).
- Numa, S. *et al. Cold Spring Harb. Symp. quant. Biol.* **48**, 57–69 (1983).
- Claudio, T. in *Frontiers in Molecular Biology: Molecular Neurobiology* (eds Glover, D. M. & Hames, B. D.) 63–142 (IRL, Oxford, 1989).
- Buonanno, A. *et al. J. Biol. Chem.* **264**, 7611–7616 (1989).
- Geisselsoder, J. *et al. Biotechniques* **5**, 786–791 (1987).
- Kunkel, T. A. *Proc. natn. Acad. Sci. U.S.A.* **82**, 488–492 (1985).
- Gu, Y. *et al. J. Cell Biol.* (in the press).

ACKNOWLEDGEMENTS. We are grateful to P. Gardner, J. Merlie, N. Davidson, J. Lindstrom and S. Froehner for cDNA clones and antibodies. We thank Y. Gu and members of Z.W.H.'s laboratory for helpful advice. This work was supported by the NIH and the Muscular Dystrophy Association.

Peptide binding to class I MHC on living cells and quantitation of complexes required for CTL lysis

E. Rosemary Christinck*, Mark A. Luscher†, Brian H. Barber† & David B. Williams*

Departments of *Biochemistry and †Immunology, University of Toronto, Toronto M5S 1A8, Canada

ANTIGENIC peptides are presented to CD8⁺ T lymphocytes by class I major histocompatibility complex (MHC) molecules^{1,2}. Peptides specifically bind to purified class I molecules *in vitro*^{3–7}, and to class I molecules on cells at nonphysiological temperatures⁸. We report here the kinetic and equilibrium parameters for the binding of radiolabelled influenza nucleoprotein peptides (NP-Y365-380 and shorter homologues) to the murine H-2D^b molecule on intact, viable cells at 37 °C. In contrast to earlier reports, we show that peptide binding is rapid and reversible, with dissociation constants ranging from nanomolar to micromolar, suggestive of typical ligand–receptor interactions. Only 10% of cell-surface D^b molecules can bind these peptides. To address the relationship between peptide binding and T-cell recognition of the antigen–MHC complex, we determined the minimum number of complexes required to sensitize a target cell for lysis by class I-restricted cytotoxic T-lymphocytes. Our data indicate that EL4 thymoma cells (H-2^b) can be sensitized for lysis by cytotoxic T-lymphocytes when as few as 200 class I–peptide complexes (less than 0.08% of surface D^b molecules) are present per cell.

A synthetic peptide corresponding to the influenza A/PR/8/1934 nucleoprotein (NP) 365–380 sequence is recognized by NP-specific D^b-restricted cytotoxic T-lymphocytes (CTL) (ref. 1), and a radiolabelled analogue has been shown recently to bind D^b *in vitro*⁶. We have employed a derivative of this peptide containing an added N-terminal tyrosine residue (NP-Y365-380) to allow labelling with ¹²⁵I. This peptide was purified by reversed-phase HPLC to exclude contaminating shorter peptides with different binding properties (ref. 9 and see below). Binding of the labelled peptide to D^b on EL4 cells was

measured by immunoprecipitation using the monoclonal antibody B22-249.R1 (B22). B22 recognizes D^b molecules associated with β_2 -microglobulin (β_2m)¹⁰. Specific binding was calculated by subtracting radioactivity immunoprecipitated from cells that had been incubated with labelled peptide in the presence of a 100-fold excess of unlabelled peptide. The overall efficiency of D^b immunoprecipitation by B22 was 74% (Table 1); all peptide binding experiments were done using a single immunoprecipitation step and the results were corrected by this efficiency factor.

The specificity of NP-Y365-380 for D^b was assessed as shown in Table 1. In a panel of four different cell lines and six anti-

TABLE 1 Specificity of NP-Y365-380 for H-2D^b

Cell type	Antibody	Class I specificity	Specific binding (c.p.m.)
EL4 ²²	B22-249.R1 ¹⁰	D ^b	12,610
	Y-3 ²³	K ^b	230
R1.1 ²⁴	15-3-1S ²⁵	D ^k , K ^k	<200
	16-1-11N ²⁵	D ^k , K ^k	<200
R1.G1 ²⁴	16-1-11N	D ^k , K ^k	240
	anti-8 ²⁶	K ^k	<200
MDAY/D2 ²⁷	34-5-8S ²⁸	D ^d	230
	anti-8	K ^d	570

Iodinated NP-Y365-380 (YIASNENMETMESSTLE) was incubated with various cell lines in the presence or absence of 100-fold excess unlabelled NP365-380. Class I molecules were immunoprecipitated using the antibodies indicated, and specific binding was calculated as the difference in radioactivity precipitated from cells when incubated in the presence of unlabelled peptide competitor versus that when incubated in its absence. Radioactivity precipitated in the presence of unlabelled peptide ranged from 1–20% of that precipitated in its absence. The data presented is representative of many similar binding experiments. In this experiment, 12,613 c.p.m. is equivalent to 7,680 D^b–peptide complexes per cell. NP-Y365-380 was radioiodinated using the chloramine-T method, and its specific activity was determined by TLC analysis of an aliquot (average 25–50 μ Ci μ g^{−1}). Unbound ¹²⁵I was removed by reversed-phase chromatography on C₁₈ Sep-Pak (Millipore). After elution, the labelled peptide was lyophilized, resuspended in Roswell Park Memorial Institute (RPMI)–BSA medium (RPMI 1640, 0.1% BSA, 2 mM glutamine), and counted to determine the peptide concentration. EL4 cells that had been grown in RPMI 1640 containing 10% fetal bovine serum were washed extensively with RPMI/BSA and then 10⁷ cells in 0.1 ml were placed in the wells of a 96-well plate. Iodinated NP-Y365-380 (1.34 μ M), unlabelled peptide (in cold blocking experiments), and protease inhibitors (1% v/v aprotinin, 0.25 mM PMSF, 0.1 mM iodoacetamide, 5 mM EDTA) were added, followed by RPMI–BSA to bring the volume to 0.15 ml. Cells were incubated at 37 °C for 2 h, washed twice at 4 °C with RPMI–BSA, lysed at 4 °C for 30 min in 0.5 ml lysis buffer (0.1 M Tris, 0.05 M NaCl, 1 mM MgCl₂, 0.5% Nonidet P-40, pH 7.4), and centrifuged for 15 min at 15,000 r.p.m. Supernatants were adjusted to 5% skimmed milk powder, and precipitated for 3 h at 4 °C with the indicated monoclonal antibody coupled to agarose. In the case of anti-8 serum, immune complexes formed during a 2 h incubation were isolated in a subsequent 1 h incubation with protein A-agarose. Agarose beads were washed four times before counting. The efficiency of precipitation and solubilization was determined as follows. (1) Radiolabelled D^b was precipitated from an EL4 cell lysate using B22. A second round of precipitation using B22 was done to estimate the residual D^b in the lysate. (2) The kinetics of the dissociation of D^b–peptide complexes in the lysate were measured at 4 °C, and the resulting first-order rate constant was used to extrapolate from the number of complexes isolated to the number present at the time of lysis. (3) The efficiency of solubilization was determined for K^b because of the availability of anti-8, a rabbit antiserum that recognizes a conformation-independent determinant in the intracellular portion of K^b (ref. 26). EL4 cells were metabolically labelled with [³⁵S]Met and, following lysis, the insoluble material was boiled in 1% SDS and diluted in phosphate-buffered saline containing 1.5% Triton X-100. Radiolabelled K^b was precipitated from this fraction with anti-8 and compared with anti-8 precipitable radioactivity recovered from the soluble portion of the original cell lysate. Taken together, the data allowed us to estimate an efficiency of 95% for a single round of B22 precipitation, 85% due to the dissociation of the class I–peptide complex, and 92% for class I solubilization, giving an overall efficiency of 74% for the isolation of D^b (data not shown). Precipitated radioactivity was divided by this efficiency factor in the calculation of class I peptide complexes per cell in all experiments.

bodies, the peptide is bound by D^b, but not by K^b, K^k, D^k, K^d or D^d. PAGE of the immunoprecipitated material under reducing conditions confirmed that precipitated radioactivity was associated with peptide and not with a transfer of radiolabel to D^b (data not shown). Thus, for this particular peptide, binding seems to be specific for the D^b molecule. In studies of the interaction of purified class I molecules with solid-phase pep-

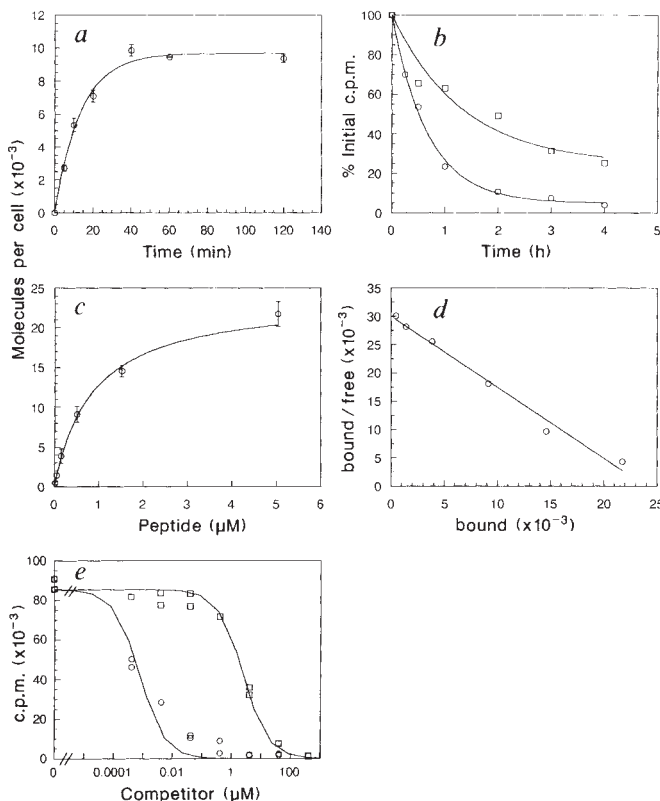
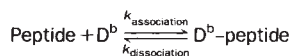


FIG. 1 Characteristics of peptide binding to H-2D^b on EL4 cells. EL4 cells were treated as in Table 1 with the modifications described, and D^b-peptide complexes were precipitated with B22. *a*, Binding kinetics: labelled NP-Y365-380 was incubated with EL4 cells for the indicated times at 37 °C under pseudo first-order conditions, that is, ~2.6 nM receptor (based on receptor number as in *d*) and 1.68 μM peptide. The observed pseudo first-order rate constant was $k_{\text{obs}} = 1.22 \pm 0.14 \times 10^{-3} \text{ s}^{-1}$ ($t_{1/2} = 9.3 \pm 1.1 \text{ min}$). *b*, Dissociation kinetics: cells were incubated with labelled NP-Y365-380 (1.68 μM) for 2 h at 37 °C, washed twice at 4 °C with RPMI-BSA, resuspended in 1 ml RPMI-BSA (plus protease inhibitors as in original incubation), and further incubated at 23 °C (□), or 37 °C (○) for various periods of time before lysis. Both reactions follow first-order kinetics: $k_{\text{dissociation}}$ (23 °C) = $1.39 \pm 0.55 \times 10^{-4} \text{ s}^{-1}$ ($t_{1/2} = 83 \pm 33 \text{ min}$); $k_{\text{dissociation}}$ (37 °C) = $4.08 \pm 0.22 \times 10^{-4} \text{ s}^{-1}$ ($t_{1/2} = 28.3 \pm 1.5 \text{ min}$). The same results were obtained if an excess of homologous unlabelled peptide was included in the incubations (data not shown). *c*, Saturation binding: cells were incubated with various concentrations of labelled NP-Y365-380 for 2 h at 37 °C. *d*, Scatchard plot of the data in *c*: maximum D^b-peptide complexes = $23,800 \pm 1,340 \text{ cell}^{-1}$ (equivalent to 30,500 c.p.m. immunoprecipitated); $K_d = 0.85 \pm 0.12 \mu\text{M}$. All curves were fit to the theoretical equation (first-order kinetics in *a* and *b*, saturation binding in *c*) using curvilinear least-squares regression analysis. Correlation coefficients were 0.99 or greater (except in *b* at 23 °C, where $r = 0.96$). Assuming a reaction of the form:



where $k_{\text{association}} = k_{\text{obs}}/1.68 \mu\text{M}$, the ratio $k_{\text{dissociation}}/k_{\text{association}}$ gives the equilibrium constant for dissociation: $K_d = 0.56 \pm 0.06 \mu\text{M}$ at 37 °C. *e*, Inhibition of labelled NP-Y365-380 binding to D^b by unlabelled NP-Y365-380 (□) or NP366-374 (○). Equilibrium binding constants were evaluated based on the known binding constant for NP-Y365-380 by curvilinear regression analysis, fitting the data to the equation for competitive binding. The resulting dissociation constant for NP366-374 was determined to be $0.46 \pm 0.25 \text{ nM}$.

tides *in vitro*, peptides that bind to one class I molecule generally bind to others^{4,5,7}. But studies on the inhibition of CTL-mediated cell lysis by diverse peptides suggest that such cross-reactivity is less common^{11,12}. Consequently, the question arises as to whether the broad cross-reactivity of peptides for class I *in vitro* truly reflects the peptide binding specificity of class I molecules or if it is an artefact of the *in vitro* assay. It should now be possible to address this issue by directly comparing binding of a variety of peptides to a panel of purified class I molecules in the *in vitro* system with the binding of the same peptides to cell-surface class I molecules using the assay we describe.

The binding of NP-Y365-380 to D^b on EL4 cells was examined as shown in Fig. 1. Binding was very rapid, with a half-time of $9.3 \pm 1.1 \text{ min}$ (peptide concentration, $1.68 \mu\text{M}$) and a dissociation half-time of $28.3 \pm 1.5 \text{ min}$ at 37 °C (Fig. 1*a, b*). This contrasts with results *in vitro* (at 20 °C)⁴ which showed a maximal binding of purified class I to solid phase peptide between 6 and 24 h of incubation as well as an extremely slow dissociation rate. The kinetics also contrast with results for peptide binding to purified class II molecules, where both peptide binding and dissociation are slow processes¹³. In intact cells, the binding of peptide to class II molecules also occurs much more rapidly¹⁴, suggesting that for both class I and class II molecules some aspect of the cellular environment may influence MHC-peptide interactions.

A Scatchard analysis revealed typical saturation binding with the maximal number of D^b-peptide complexes being $23,800 \pm 1,340$ (Fig. 1*c, d*). Binding was not significantly altered by pre-treatment of the cells for 2 h with the protein synthesis inhibitor cycloheximide, suggesting that binding occurs predominantly to cell-surface localized class I molecules (data not shown). Using radiolabelled F(ab) fragments of B22, the surface-accessible pool of D^b molecules was determined to be 258,000 per EL4 cell (data not shown). Thus, only a small amount (9.2%) of the total accessible surface D^b molecules was able to bind peptide. Whether this small fraction of D^b is able to accept peptide by virtue of being 'empty', or being in possession of a lower affinity peptide that can be displaced by exchange, remains to be determined. Exogenous β₂m promotes the formation of class I-peptide complexes either with intact cells¹⁵⁻¹⁸ or with purified class I molecules¹⁸, suggesting that peptide binding may occur during β₂m exchange. This is unlikely to be the mechanism for the binding of peptide to D^b here because β₂m was excluded from the incubation medium. It will be of interest to determine if the fraction of D^b molecules that binds peptide can be increased by the addition of exogenous β₂m in our assay.

We estimated the dissociation constant for the reaction by two independent methods, first from the saturation binding data (dissociation constant $K_d = 0.85 \pm 0.12 \mu\text{M}$), and second as a ratio of reverse to forward rate constants ($K_d = 0.55 \pm 0.06 \mu\text{M}$). Again, these results differ from those obtained *in vitro*. Frelinger *et al.*⁵ observed a K_d of about 600 μM for the interaction of detergent-solubilized HLA-A2 with the peptide KGILGFVFKLTV (single-letter amino-acid code) bound to polyvinylchloride plates. Because of the potential for effective multivalency of both the ligand (plate-bound peptide) and the receptor (class I in detergent micelles), such binding data are difficult to interpret. But in the cell-binding assay described here, peptide is free in solution and thus the dissociation constant and rate data should reflect the interaction of a single peptide with a single class I molecule in the absence of such considerations. Nevertheless, the data obtained with this assay may be specific to the D^b interaction with NP-Y365-380 and not characteristic of all class I-peptide interactions.

Recent studies have identified naturally processed viral peptides recognized by CTL^{9,19}. In the case of D^b-restricted recognition of influenza nucleoprotein, the natural peptide was determined to be NP366-374. It was recognized by CTL much more effectively than NP365-380 (ref. 9). To understand the basis for the increased potency of the natural peptide, we synthesized this peptide with an N-terminal tyrosine (NP-Y366-374),

purified it by HPLC, and examined the binding of the radioiodinated peptide to D^b on EL4 cells. Saturation binding studies revealed that whereas the maximum number of D^b-peptide complexes obtained for NP-Y366-374 was comparable to that for NP-Y365-380, the dissociation constant for the shorter peptide was about 100-fold lower ($K_d = 7.7 \pm 3.8$ nM versus ~ 700 nM; data not shown). To assess the binding of the natural peptide without the N-terminal tyrosine, we synthesized and purified NP366-374 and tested its ability to inhibit the binding of radioiodinated NP-Y365-380 to D^b (Fig. 1e). NP366-374 was a much more potent inhibitor of binding compared with unlabelled NP-Y365-380. The dissociation constant calculated for the natural peptide ($K_d = 0.46 \pm 0.25$ nM) was 1,000–2,000-fold lower than that for NP-Y365-380. Thus, the very efficient recognition of the naturally processed peptide by CTL can be attributed largely to higher affinity binding of the peptide to the D^b molecule. Furthermore, the addition of a single tyrosine to the N terminus of this peptide substantially reduces the binding affinity.

Having established a quantitative basis for determining the number of D^b-peptide complexes on the surface of intact cells at different peptide concentrations, we applied this methodology to address the question of how many class I-peptide complexes are required for effective recognition and lysis by CTL. We used an uncloned, X-31-specific CTL line, NPL.16, that recognizes

NP-Y365-380 in association with D^b. Figure 2 is a representative example of three experiments in which the lysis of EL4 cells by NPL.16 was examined at various peptide concentrations. Sub-optimal lysis was observed at 0.025 μ M peptide, a concentration where as few as 200 ± 24 D^b-bound peptides (less than 0.08% of total surface D^b) were present on the surface of each target cell. A comparison of Fig. 2a and b shows that lysis of cells incubated with radiolabelled peptide was not significantly different from the lysis of cells incubated with the native peptide. This suggests that the radiolabelled peptide binds D^b with a similar affinity to unlabelled peptide, is recognized equally well by the CTL effector, and is not significantly degraded in the labelling procedure.

We repeated the experiment with the radiolabelled derivative of the naturally processed peptide (NP-Y366-374). As with the longer peptide, lysis could be detected with as few as 200 D^b-peptide complexes per cell. But as expected from the higher binding affinity, this number of complexes was created at a substantially lower peptide concentration (0.0005 μ M; data not shown). Although the number of D^b-peptide complexes required for lysis is low, we believe that this figure represents an upper limit. For example, it is not known if all peptide molecules are bound in a conformation relevant for recognition by the T-cell receptor, and how many of these roughly 200 complexes initially distributed over the target cell surface are collected into the point of productive interaction with the CTL. In addition, whereas about 200 complexes are found on the cell surface at the initiation of the 4 h CTL assay, our kinetic data show that this number decreases with a half-time of 28 min as the assay progresses, a factor that further diminishes the number of available target structures. Recently, an indirect flow cytometric approach was used to estimate an upper limit for the number of complexes required for lysis by CTL at about 200–400, consistent with our findings¹⁷.

The sensitivity of the T cell in recognizing its antigen should be dictated in part by the affinity of its T-cell receptor. Thus we used a bulk CTL line as the effector population to represent an 'average' affinity, rather than a cloned T cell which might have an unusually high or low affinity. Our data, however, are in good agreement with recent results showing that 60–300 peptides are necessary to stimulate interleukin-2 production in class II-restricted T-cell hybridomas^{20,21}, suggesting that these hybridomas are of a similar affinity to the T-cell line used in this study. The concordance between the results obtained in the class I and class II systems also supports the concept that the recognition by T cells of antigen in association with both MHC class I and class II molecules is a highly analogous event.

Note added in proof: In binding experiments with the non-physiological peptide NP-Y365-380, we confirmed by HPLC analysis that the peptide eluted from H-2D^b immunoprecipitates exhibited the chromatographic properties of NP-Y365-380 rather than those of shorter, homologous peptides. □

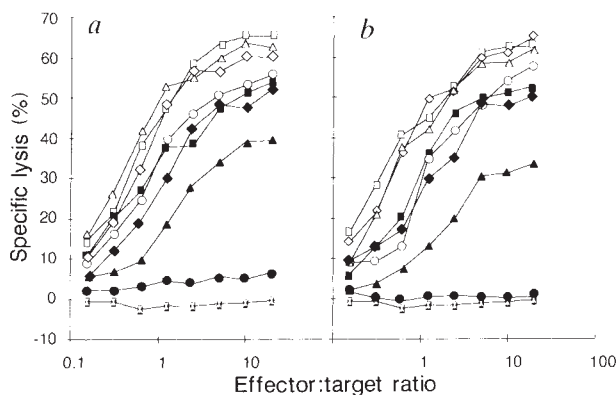


FIG. 2 Quantitation of D^b-peptide target structures required to sensitize EL4 cells for lysis by an influenza A/X-31-specific CTL line, NPL.16. EL4 cells were incubated for 2 h with various concentrations of either a, unlabelled or b, ¹²⁵I-labelled NP-Y365-380. Both the lysis by NPL.16 and the number of D^b-peptide complexes per cell were determined at the following peptide concentrations, expressed as μ M (D^b-peptide complexes per cell in brackets): □, 5.0 (5,810); ◇, 0.80 (3,360); △, 0.40 (2,240); ○, 0.20 (1,340); ■, 0.10 (740); ◆, 0.050 (390); ▲, 0.025 (200); ●, 0.0050 (40); ▨, 0 (–). **METHODS.** NPL.16 was derived from spleen cells of C57BL/6J mice (Jackson Laboratories, Bar Harbor, Maine) 11 weeks after infection with 100 haemagglutinating units of influenza virus strain X-31. X-31 is a recombinant influenza virus that possesses the nucleoprotein of influenza A/PR/8/1934 (ref. 29). Cells were maintained in culture by weekly passage on irradiated (2,000R) X-31-infected spleen cells with the addition of medium containing interleukin-2 after the second passage. NPL.16 recognized NP-Y365-380 restricted to D^b (data not shown), and was used 4–8 weeks after initiation. For the CTL assay, EL4 cells were incubated with peptide at various concentrations as described in Table 1, except that 150 μ Cl Na₂⁵⁵CrO₄ per 10⁷ cells were also added. Cells were washed extensively at 22 °C, and added to titrated effectors for a 4-h ⁵⁵Cr-release assay. Specific release is reported as the mean of triplicate values. The number of D^b-bound peptides per cell was determined by immunoprecipitation using B22 as in Table 1. Cell lysates for immunoprecipitation were prepared at the time of initiation of the CTL assay, from aliquots of the EL4 target cells used in the CTL assay. A saturation curve was prepared as in Fig. 1c, and D^b-bound peptides per cell were estimated by interpolation on the best-fit curve (not shown; correlation coefficient = 0.996). Error in the determination is estimated as 12% of the value given (for example, ▲, 200 ± 24 peptides per cell). Note that the numbers of complexes per cell in this experiment are reduced relative to Fig. 1c. This is due to complex dissociation during the extensive washing (~ 2 h) that preceded the CTL and immunoprecipitation assays.

Received 25 April; accepted 17 May 1991.

1. Townsend, A. R. M. *et al. Cell* **44**, 959–968 (1986).
2. Townsend, A. & Bodmer, H. A. *Rev. Immun.* **7**, 601–624 (1989).
3. Chen, B. P. & Parham, P. *Nature* **337**, 743–745 (1989).
4. Chopin, J. *et al. J. exp. Med.* **172**, 889–899 (1990).
5. Frelinger, J. A., Gotch, M. M., Zweerink, H., Wain, E. & McMichael, A. J. *J. exp. Med.* **172**, 827–834 (1990).
6. Townsend, A. *et al. Cell* **62**, 285–295 (1990).
7. Chen, B. P., Rothbard, J. & Parham, P. *J. exp. Med.* **172**, 931–936.
8. Schumacher, T. N. M. *et al. Cell* **62**, 563–567 (1990).
9. Rötzschke, O. *et al. Nature* **348**, 252–254 (1990).
10. Lemke, H., Hämmerling, G. & Hämmerling, U. *Immunol. Rev.* **47**, 175–206 (1976).
11. Carreno, B. M., Anderson, R. W., Coligan, J. E. & Biddison, W. E. *Proc. natn. Acad. Sci. U.S.A.* **87**, 3420–3424 (1990).
12. Bodmer, H. C., Bastin, J. M., Askonas, B. A. & Townsend, A. R. M. *Immunology* **66**, 163–169 (1989).
13. Buus, S., Sette, A., Colon, S. M., Jenis, D. M. & Grey, H. M. *Cell* **47**, 1071–1077 (1986).
14. Ceppellini, R. *et al. Nature* **339**, 392–394 (1989).
15. Rock, K. L., Rothstein, L. E., Gamble, S. R. & Benacerraf, B. *Proc. natn. Acad. Sci. U.S.A.* **87**, 7517–7521 (1990).
16. Rock, K. L., Gamble, S., Rothstein, L. & Benacerraf, B. *Proc. natn. Acad. Sci. U.S.A.* **88**, 301–304 (1991).
17. Vitiello, A., Potter, T. A. & Sherman, L. A. *Science* **250**, 1423–1426 (1990).

18. Kozlowski, S. et al. *Nature* **349**, 74–77 (1991).
19. Van Bleek, G. M. & Nathanson, S. G. *Nature* **348**, 213–216 (1990).
20. Harding, C. V. & Unanue, E. R. *Nature* **346**, 574–576 (1990).
21. Demotz, S., Grey, H. & Sette, A. *Science* **249**, 1028–30 (1990).
22. Gorer, P. A. *Br. J. Cancer* **4**, 372–380 (1950).
23. Hämmerling, G. J., Rüsche, E., Tada, N., Kimura, S. & Hämmerling, U. *Proc. natn. Acad. Sci. U.S.A.* **79**, 4737–4741 (1982).
24. Hyman, R. & Stallings, V. *Immunogenetics* **3**, 75–84 (1976).
25. Ozato, K., Mayer, N. & Sachs, D. H. *J. Immun.* **124**, 533–540 (1980).
26. Smith, M. H., Parker, J. R. M., Hodges, R. S. & Barber, B. H. *Molec. Immun.* **23**, 1077–1092 (1986).
27. Lagarde, A. E., Donaghy, T. P., Kerbel, R. & Siminovich, L. *Somatic Cell molec. Genet.* **10**, 503–519 (1984).
28. Ozato, K., Mayer, N. M. & Sachs, D. H. *Transplantation* **34**, (1982).
29. Kilbourne, E. D. *Bull. Wild Hlth Org.* **41**, 643–645 (1969).

ACKNOWLEDGEMENTS. E.R.C. and M.A.L. are joint first authors. We thank P. Chong of Connaught Laboratories for synthetic peptides and E. Degen for reading of the manuscript. This work was supported by the MRC of Canada (B.H.B.) and the National Cancer Institute of Canada (D.B.W.). M.A.L. and D.B.W. are a Student and Scholar, respectively, of the MRC of Canada.

Defective acidification of intracellular organelles in cystic fibrosis

Jonathan Barasch, Brian Kiss, Alice Prince, Lisa Saiman, Dieter Gruenert* & Qais Al-Awqati

Departments of Medicine, Physiology and Pediatrics, College of Physicians and Surgeons of Columbia University, New York 10032, USA

* Cystic Fibrosis Research Center, University of California, San Francisco, California 94143, USA

THE phenotype of cystic fibrosis (CF) includes abnormalities in transepithelial transport of Cl^- (refs 1–5), decreased sialylation and increased sulphation and fucosylation of glycoproteins^{6–9}, and lung colonization with *Pseudomonas*. It is not apparent how these abnormalities are interrelated, nor how they result from loss of function of the CF gene-encoded transmembrane regulator (CFTR)¹⁰. We have previously shown that the pH of a secretory granule is regulated by the vesicular conductance for Cl^- (ref. 11). Here we find defective acidification in CF cells of the *trans*-Golgi/*trans*-Golgi network, of prelysosomes and of endosomes as a result of diminished Cl^- conductance. Sialylation of proteins and lipids is reduced and ligand traffic altered. These abnormalities can result from defective acidification because vacuolar pH regulates glycoprotein processing and ligand transport. The CF phenotype is similar to that of alkalinized cells¹² and acidification-defective mutants¹³.

To analyse the transmembrane pH gradient (ΔpH) of individual organelles, we incubated nasal polyps from children with CF or allergic rhinitis with the weak base DAMP (3-(2,4-dinitro anilino)-3' amino-*N*-methylpropylamine)¹⁴. Lysosomes of epithelia from both took up DAMP, as did the *trans*-Golgi/network and the prelysosomal body¹⁵ (Fig. 1a, c) of normal cells. In contrast, CF Golgi and prelysosomes did not trap DAMP (Fig. 1b, d). Using simian virus-40 (SV40)-immortalized epithelium from CF nasal polyp and from normal trachea¹⁶, we identified the organelles by labelling simultaneously with DAMP and an antibody against the mannose-6-phosphate receptor (M6PR) (Table 1); in three experiments we found that the pH of CF organelles was always statistically significantly higher than those of the normal cells.

We also directly measured the pH of endosomes with fluorescein-tagged transferrin (Tf-F)¹⁷. Immortalized cells (10^7 , maintained in serum-free medium for 8 hours) were incubated with Tf-F ($20 \mu\text{g ml}^{-1}$) for 15 min at 37°C , washed and analysed by ratio spectrofluorometry (490 nm/460 nm excitation; 510 nm emission), standardized with formaldehyde-fixed, Tf-F-loaded cells¹⁷. We found that the pH of CF endosomes was 6.8 and that of normal endosomes was 6.3 (fluorescence ratio $\pm$ s.e.m.: CF = 1.5 ± 0.03 ; normal = 1.35 ± 0.02 ; $n = 53$, $P < 0.001$).

Ammonium chloride (50 mM) raised the fluorescence ratio by 50–200%, indicating that Tf-F was a good measure of organellar pH. These data indicate that endosomes, like other organelles in cells from CF patients, are defective in generating an acidic microenvironment.

Proton-translocating ATPases generate a ΔpH and a transmembrane potential, $\Delta\psi$, which are inversely related in the extent of their formation¹⁸. To study these gradients we isolated a 'light' and 'heavy' vesicle fraction. Light vesicles from normal cells rapidly trapped acridine orange in response to ATP (Fig. 2a), reflecting ΔpH formation. A transmembrane potential was not generated because acidification was not stimulated by the electrogenic K^+ ionophore valinomycin. This implies that the conductance of these membranes for counterions is high. In fact, acidification depends on a conductance for Cl^- rather than K^+ as dye trapping was inhibited when potassium gluconate replaced KCl (Fig. 2c). In contrast, acidification of CF light vesicles was slower and was stimulated by valinomycin (Fig. 2b). This indicates that ΔpH was limited by $\Delta\psi$, which must result from a low endogenous membrane conductance for Cl^- . CF and control heavy vesicles (enriched in lysosomes) acidified equally and were minimally stimulated by valinomycin (Fig. 2d, e). There was also no reduction in ΔpH when Cl^- was replaced with gluconate (Fig. 2f), suggesting that CF heavy vesicles have a conductance for K^+ . These data indicate that limited acidification of Golgi vesicles and endosomes in CF cells (light vesicles) is probably due to defective Cl^- conductance, and that the nearly equal acidification of lysosomes results from an unaffected K^+ conductance (heavy vesicles).

Sialyltransferases have acidic pH optima^{19–21} and are resident in the *trans*-Golgi/network²². We assayed protein sialylation by incubating immortalized cells overnight with [^3H]N-acetyl mannosamine (30 μCi), the precursor of sialic acid²³, and desialylated transferrin¹³ (at $20 \mu\text{g ml}^{-1}$ for 37°C) for 4 hours. Transferrin, recovered from the supernatant of normal cells by trichloroacetic acid (TCA) precipitation and SDS-PAGE, had a twofold greater incorporation of neuraminidase-sensitive¹³ sialic acid than did transferrin recovered from CF cells (data not shown). The cells incorporate equal quantities of *N*-acetyl mannosamine and asialotransferrin, the latter determined using

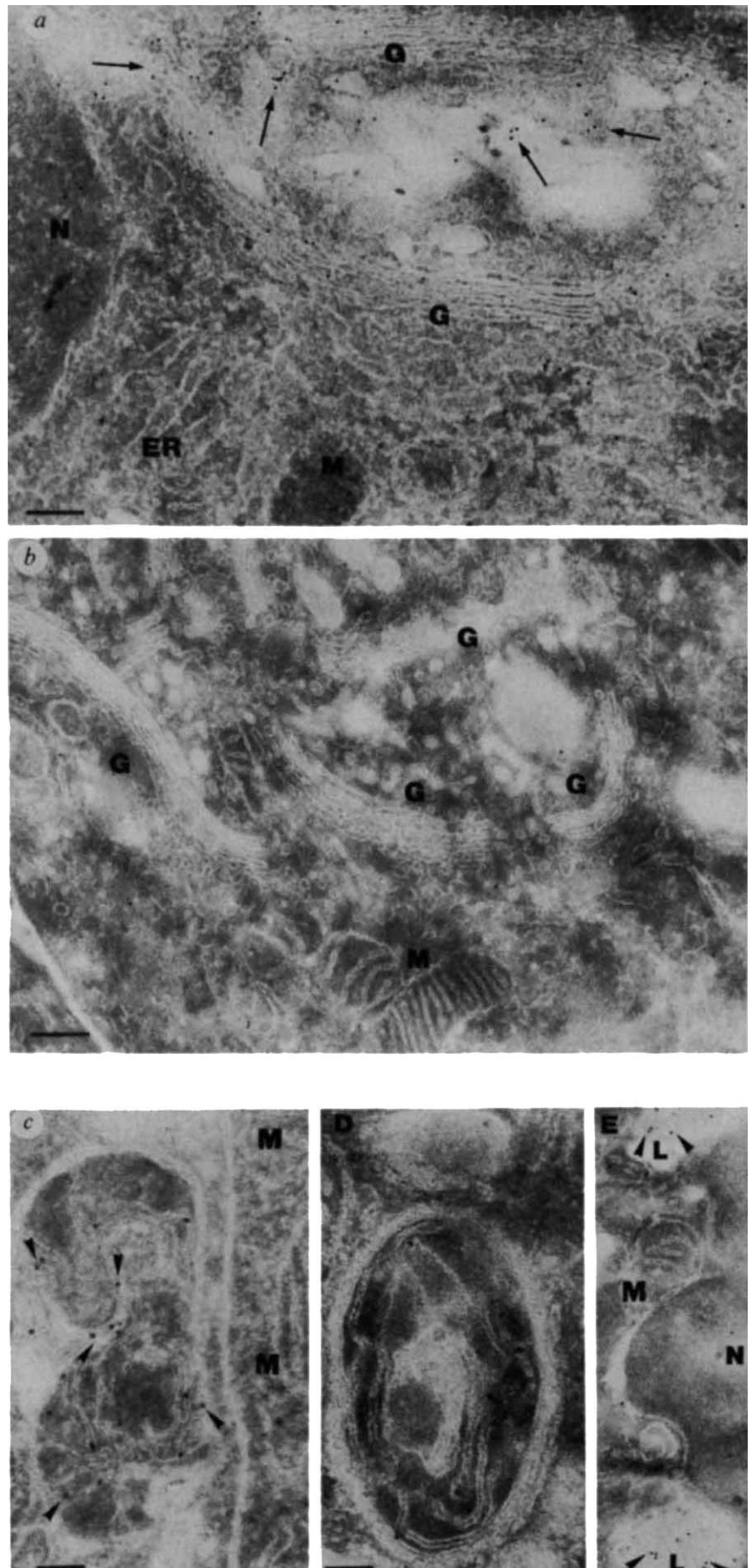
TABLE 1 Quantitative analysis of DAMP accumulation in organelles of SV40-immortalized epithelial cells

Organelle	Gold particles per μm^2		[DAMP] in organelle/[DAMP] in nucleus		pH	
	Normal	CF	Normal	CF	Normal	CF
Nucleus	10 ± 0.7	10 ± 0.5	1	1	7	7
Cis-Golgi	8 ± 1	11 ± 2	1.1 ± 0.2	1.2 ± 0.3	7	6.9
Trans-Golgi*	26 ± 3	16 ± 3	3.4 ± 0.3	1.7 ± 0.3	6.5	6.8
Pre-lysosome*	67 ± 7	38 ± 3	6.9 ± 0.7	4.4 ± 0.5	6.2	6.4
Lysosome	164 ± 21	200 ± 19	25 ± 4	31 ± 5	5.6	5.5

Cells were treated and processed as described for Fig. 1. Organelles were tentatively identified by their morphology and by their ability to generate a transmembrane pH gradient. *Trans*-Golgi/network was identified at one face of the Golgi stack by a concentration of DAMP^{16,17,36} not present at the opposing face (*cis*-Golgi). Prelysosomes were identified by their unique morphology—large, multiseptated electron-dense bodies, which simultaneously stained for DAMP and the mannose 6-phosphate receptor. Lysosomes, in contrast, were homogeneous electron-dense organelles, without internal compartmentalization, and were negative for the mannose 6-phosphate receptor and positive for DAMP. A subset of organelles were found that concentrated DAMP heavily but were also positive for M6PR. These organelles had no internal membrane structures and were of the same size and acidity as the lysosomes; they could be transitional structures. Gold particles were counted per μm^2 of nucleus, prelysosome and lysosome and in a ribbon of cytoplasm along the length of each Golgi face, which was 200 nm wide. Gold particles per μm^2 of nucleus were equated with a pH of 7, and the 'apparent pH' of each organelle determined by $\text{pH} = 7 - \log([\text{DAMP}] \text{ in organelle}/[\text{DAMP}] \text{ in nucleus})$. Two other similar experiments were performed; although the DAMP concentration per organelle varied in each experiment, the pH of CF organelles was always statistically significantly higher than the normal, albeit to different extents.

* Normal and CF significantly different ($P < 0.001$); [DAMP] $\pm$ s.e.m.

FIG. 1 Electron microscopic immunocytochemistry of normal (a, c) and CF (b, d, e) nasal polyp epithelial cells incubated with DAMP (30 μ M) for 30 min and then fixed (in 4% paraformaldehyde, 1% glutaraldehyde, 3% sucrose in 0.1M sodium phosphate, pH 7.4, at 37 °C for 3 h) and embedded in LR White¹¹. DAMP is concentrated in normal *trans*-Golgi/*trans*-Golgi network (a) and prelysosomes (c) and CF lysosomes (e), but not in CF Golgi (b) or CF prelysosomes (d). DAMP was visualized in thin sections on Formvar-coated nickel grids with anti-dinitrophenol (Oxford) in 4% horse serum in PBS at a 1:20 dilution, and the primary antibody was located with 10-nm colloidal gold-coupled anti-mouse antibody (Amersham). Organelles were tentatively identified morphologically and by demonstration of a Δ pH. The *trans*-Golgi/network is identified by DAMP labelling of small vesicles adjacent to the Golgi stacks (DAMP is a *trans*-Golgi/network marker^{14,36}); the *cis*-Golgi face shows no DAMP labelling. Pelysosomes are distinguished from lysosomes by their larger size and the presence of multiple internal membranes, and were independently identified by staining for the mannose 6-phosphate receptor (relative molecular mass, M_r , 214,000; gift from S. Kornfeld). Scale, 200 nm (a, b); 100 nm (c, d); 250 nm (e). G, Golgi; N, nucleus; M, mitochondria; ER, endoplasmic reticulum, L, lysosome. The apparent pHs of CF *trans*-Golgi network and prelysosomes are alkalinized by 0.4 and 0.25 pH units respectively; lysosomes are not affected (see Table 1 legend for method of quantitation).



^{125}I -labelled asialo-transferrin, which demonstrates that resialylation is reduced in CF. To determine whether sialylation of all glycoproteins is abnormal in CF, we incubated the cells with [^3H]N-acetyl mannosamine and [^{35}S]methionine. The secreted proteins gave similar [^{35}S]methionine patterns on SDS-PAGE (Fig. 3a), but proteins from normal cells were more heavily sialylated (Fig. 3b) per methionine residue than CF proteins (Fig. 3c, d). Thus, although protein synthesis and secretion is not generally altered, terminal glycosylation of proteins in the *trans*-Golgi/network is defective in CF. We also metabolically labelled lipids by incubating cells for three days with [^{14}C]butyrate (6 μCi) and N-acetyl mannosamine (5mM). The latter was included so that synthesis of sialo-glycolipids would not be limited by precursor availability. Cells were then extracted²⁴, and GM₁ and asialo-GM₁ gangliosides analysed with thin-layer chromatography (TLC)²⁴ by spotting equal [^{14}C] counts and recovering the gangliosides comigrating with standards. CF cells had more asialo-GM₁ than did control cells (CF, 18%; control, 14% of lipid spotted on TLC) and a higher asialo-GM₁/GM₁ ratio (CF, 2.8; control, 2.3; $n = 2$). Although SV40 transfection itself decreases GM₁ levels²⁵, the data suggest that sialylation of both glycolipids and glycoproteins is reduced in CF. As both classes of sialyltransferases have acidic pH optima, reduced sialylation probably results from alkalinization of the *trans*-Golgi/network.

The routing of soluble proteins through cells depends on ligand-receptor interactions that are pH-sensitive¹². When we

loaded endosomes with [^{125}I] α_2 -macroglobulin¹⁷, both cell types took up equal amounts of protein and TCA-soluble peptides appeared rapidly from normal cells. But degradation was consistently delayed by >10 min in CF cells (Fig. 4). We propose that alkalinization of the CF endosomal pathway either reduced ligand-receptor uncoupling or alters interorganellar vesicle movement, delaying ligand transfer from endosomes to lysosomes. But as lysosomal pH is not altered in CF, the rate of proteolysis after ligand delivery is similar to the control cells.

The routing of transferrin in CF was analysed by incubating cells with 20-nm colloidal gold coupled to transferrin²⁶ (20 μg protein per ml; incubated for 4 h at 37 °C) and locating the probe in the endocytic pathway by simultaneous electron microscope immunocytochemistry for the mannose 6-phosphate receptor and for DAMP (see Fig. 1 legend). In normal cells ($n = 127$), 9% of the probe was found in M6PR-negative, DAMP-positive lysosomes (68% in M6PR and DAMP-negative clear endosomes, 22% in M6PR and DAMP-positive prelysosomes), whereas in CF cells ($n = 141$), 23% was in lysosomes (65% in endosomes, 12% in prelysosomes). Preincubation with a tenfold excess of unlabelled transferrin abolished uptake of the probe, indicating that the transferrin receptor was responsible for transport. In the acid pH of normal endosomes/prelysosomes, iron dissociation generates apo-transferrin-gold, which recycles to the cell surface by tightly binding its receptor²⁷. But in the less acidic CF endosomes, dissociation of apo-transferrin-gold from its receptor allows transfer to lysosomes.

As membrane Cl^- conductance regulates vesicular pH^{11,12,18,28,29}, and acidification of the light-vesicle fraction depends on Cl^- , decreased acidification of CF organelles probably results from defective Cl^- conductance. Thus CF manifests with defective chloride transport both on the plasma membrane and in intracellular organelles. This may result from the absence of the CFTR in the post-endoplasmic reticulum vacuolar system and subsequent loss of channel regulation in Golgi, prelyso-

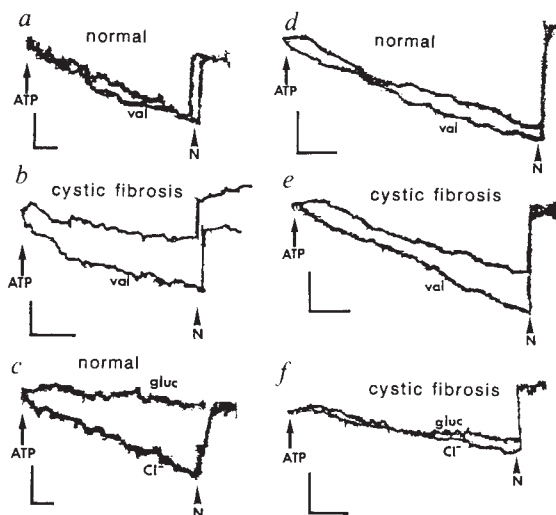


FIG. 2 Acidification of light and heavy vesicles assayed spectrophotometrically by uptake of acridine orange (490 nm excitation, 530 nm emission). Light vesicles (a, b, c) or heavy vesicles (d, e, f) (200 μg protein in 150 mM KCl, 6 mM MgSO_4 , 2 mM Tris/MES, pH 7, plus 1 mM dithiothreitol, acridine orange (6 $\mu\text{g ml}^{-1}$), oligomycin (50 $\mu\text{g ml}^{-1}$), and 1 μM valinomycin (where indicated by 'val')) were treated with Tris-ATP (0.3 mM); ΔpH formation was detected by a downward deflection. Concentration of the dye is reversed by K^+/H^+ exchange with nigericin ('N'; 2 μM), indicating that dye trapping reflected ΔpH formation. Chloride ion-free buffer is the same as above, except that it contains 150 mM potassium gluconate and gluconic acid is used for titration of pH (c, e, 'gluc'). Scale bars for x axis, 2 min; for y axis, 5% of scale. To prepare vesicles, immortalized CF and normal cells were incubated in serum-free minimal essential medium (MEM) overnight, homogenized with a Dounce homogenizer in 0.3M sucrose, 3 mM HEPES and 1 mM EDTA, pH 7.2, and a postnuclear supernatant prepared (spun at 1,200g for 10 min). Supernatant (15 ml) was layered over a discontinuous metrizamide gradient¹¹, centrifuged at 10,000g for 45 min, and the supernatant collected as 'light' vesicles. Cells were also preincubated with [^{125}I]labelled asialo-transferrin (1.8 μCi ; 0.7 $\mu\text{Ci per } \mu\text{g}$) for 1 h, washed for 30 min, and then homogenized. 'Heavy' vesicles were prepared from the postnuclear supernatant on a linear sucrose gradient (0.5–1.3M sucrose with 3 mM HEPES; spun at 100,000g for 90 min) in a dense fraction enriched in acid phosphatase activity and concentrating [^{125}I]labelled asialo-transferrin (30–45% recovered counts).

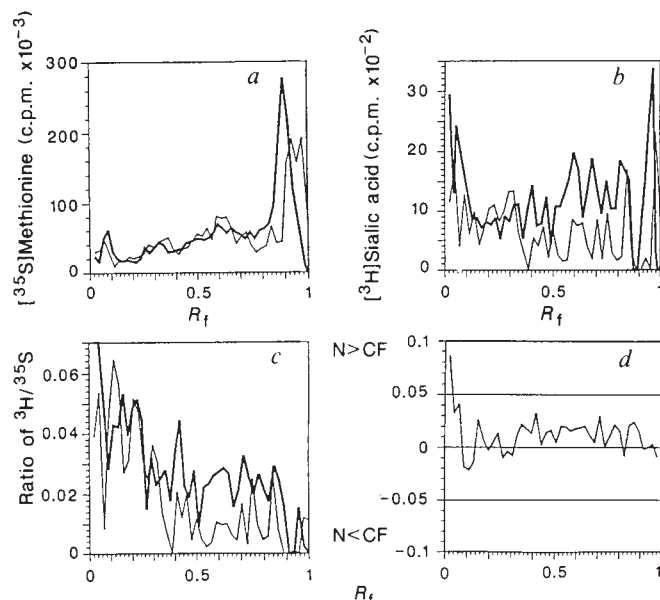


FIG. 3 Sialylation of proteins secreted from immortalized cell lines was detected by incubating cells in serum-free MEM with [^3H]N-acetyl mannosamine (30 μCi) and [^{35}S]methionine (138 μCi) for 18 h, and analysing the cell supernatant by precipitation with trichloroacetic acid (TCA) (8%) and SDS-PAGE (6%). The [^{35}S]methionine patterns are nearly identical (a), but sialylation was greater in normal cell supernatants (bold line) than in those from CF cells (thin line) (b). Correcting for protein concentration in each gel slice ($^3\text{H}/^{35}\text{S}$) confirms that there is more sialylation of normal than of CF glycoproteins (c). d, Difference between normal (N) and CF sialic acid/methionine ratios, showing more extensive sialylation of proteins from normal cells across an M_r range of 200–30,000.

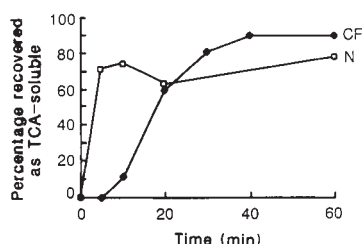


FIG. 4 Proteolysis of α_2 -macroglobulin. Time course of the delivery of ^{125}I -labelled α_2 -macroglobulin to lysosomes is assayed by the appearance of TCA-soluble ^{125}I -labelled peptides. Appearance of degraded protein is rapid in normal cells (open squares), but delayed in CF cells by >10 min (filled diamonds). Cells (in serum-free medium, incubated overnight) were incubated with 0.1 μCi (18 μg) protein for 15 min at 37 °C (plus 5 min for medium to warm up), washed and reincubated in MEM (ref. 17) (initially at room temperature). Counts are expressed as a percentage of total counts per dish. The lengthy preincubation with the probe was to ensure loading of the endocytic pathway to allow assay of endosome-lysosome transfer of the probe. α_2 -Macroglobulin was a gift from F. Maxfield.

somes and endosomes, as the mutant protein is confined to a perinuclear distribution³⁰. Immature glycosylation of CFTR³⁰, then, is secondary to a failure of delivery to the Golgi, rather than a consequence of abnormal late Golgi and vacuolar pH.

Consequences of vacuolar alkalization may include (1) decreased protein sialylation and increased fucosylation and sulphation^{31,32}, as the latter are competitive with sialylation and have neutral pH optima; (2) increased asialo-gangliosides, which are putative *Pseudomonas* binding sites³³ and would therefore explain the predominance of this organism in CF; (3) inappropriate secretion of lysosomal enzymes, exacerbating airway dysfunction; (4) abnormal glycosylation of mucus glycoproteins, which could alter their rheological properties, and (5) heterozygote advantage as a result of reducing sialylated GM₁ cholera receptors³⁴ and limiting pH-dependent virus and toxin penetration of endosomes³⁵. □

Received 27 December 1990; accepted 9 May 1991.

1. Quinton, P. M. *Nature* **301**, 421–422 (1983).
2. Knowles, M. R. *et al. Science* **221**, 1067–1069 (1983).
3. Schoumacher, R. A. *et al. Nature* **330**, 752–754 (1987).
4. Li, M. *et al. Nature* **331**, 358–360 (1988).
5. Hwang, T. C. *et al. Science* **244**, 1351–1353 (1989).
6. Boat, T. F. *et al. Am. Rev. resp. Dis.* **110**, 428–441 (1974).
7. Frates, R. C., Kaizu, T. T. & Last, J. A. *Pediatr. Res.* **17**, 30–34 (1983).
8. Wesley, A., Forstner, J., Qureshi, R., Mantle, M. & Forstner, G. *Pediatr. Res.* **17**, 65–69 (1983).
9. Scanlin, T. F., Wang, Y.-M. & Glick, M. C. *Pediatr. Res.* **19**, 368–373 (1985).
10. Riordan, J. R. *et al. Science* **245**, 1066–1073 (1989).
11. Barasch, J., Gershon, M. D., Nunez, E. A., Tamir, H. & Al-Awqati, Q. *J. Cell Biol.* **107**, 2137–2147 (1988).
12. Mellman, I., Fuchs, R. & Helenius, A. A. *Rev. Biochem.* **55**, 663–700 (1986).
13. Roff, C. F., Fuchs, R., Mellman, I. & Robbins, A. R. *J. Cell Biol.* **103**, 2283–2297 (1986).
14. Anderson, R. G. W. & Pathak, R. *Cell* **40**, 635–643 (1985).
15. Griffiths, G., Holflack, B., Simons, K., Mellman, I. & Kornfeld, S. *Cell* **52**, 329–341 (1988).
16. Wagner, J. A. *et al. Nature* **349**, 793–796 (1991).
17. Yamashiro, D. J., Tycko, B., Fluss, S. R. & Maxfield, F. R. *Cell* **37**, 789–800 (1984).
18. Al-Awqati, Q. A. *Rev. Cell Biol.* **2**, 179–199 (1986).
19. Pohlentz, G., Klein, D., Schwarzmann, G., Schmitz, D. & Sandhoff, K. *Proc. natn. Acad. Sci.* **85**, 7044–7048 (1988).
20. Busam, K. & Decker, K. *Eur. J. Biochem.* **160**, 23–30 (1986).
21. Roseman, S. *et al. Meth. Enzym.* **8**, 354–372 (1966).
22. Taatjes, D. J. & Roth, J. *Eur. J. Cell Biol.* **42**, 344–350 (1986).
23. Rosner, H., Wiegandt, H. & Rahmann, H. *J. Neurochem.* **21**, 655–665 (1973).
24. Svennerholm, L. & Fredman, P. *Biochim. Biophys. Acta* **617**, 97–109 (1980).
25. Fishman, P. H. & Brady, R. O. *Science* **194**, 906–915 (1976).
26. Dickson, R. B., Willingham, M. C. & Pastan, I. *J. Cell Biol.* **89**, 29–34 (1981).
27. Klausner, R. D. *et al. Proc. natn. Acad. Sci. U.S.A.* **80**, 2263–2266 (1983).
28. Johnson, R. G. & Scarpa, A. *J. Biol. Chem.* **254**, 3750–3760 (1979).
29. Bae, H. & Verkman, A. S. *Nature* **348**, 637–639 (1990).
30. Cheng, S. H. *et al. Cell* **63**, 827–834 (1990).
31. Paulson, J. C., Prieels, J.-P., Glasgow, L. R. & Hill, R. L. *J. Biol. Chem.* **253**, 5617–5624 (1978).
32. Green, E. D., Gruenbaum, J., Bielinska, M., Baenziger, J. U. & Boime, I. *Proc. natn. Acad. Sci. U.S.A.* **81**, 5320–5324 (1984).
33. Kriven, H. C., Ginsburg, V. & Roberts, D. D. *Arch. Biochem. Biophys.* **260**, 493–496 (1988).
34. Fishman, P. H., Moss, J. & Vaughan, M. *J. Biol. Chem.* **251**, 4490–4494 (1976).
35. White, J., Kielian, M. & Helenius, A. *Q. Rev. Biophys.* **16**, 151–195 (1983).
36. Griffiths, G. & Simons, K. *Science* **234**, 438–443 (1986).

ACKNOWLEDGEMENTS. We thank S. Kornfeld for the mannose 6-phosphate receptor and F. Maxfield for α_2 -macroglobulin. This work was supported by the US National Institutes of Health and the US Cystic Fibrosis Foundation.

Structure of the insulin receptor substrate IRS-1 defines a unique signal transduction protein

Xiao Jian Sun, Paul Rothenberg*, C. Ronald Kahn, Jonathan M. Backer, Eiichi Araki, Peter A. Wilden, Deborah A. Cahill, Barry J. Goldstein & Morris F. White

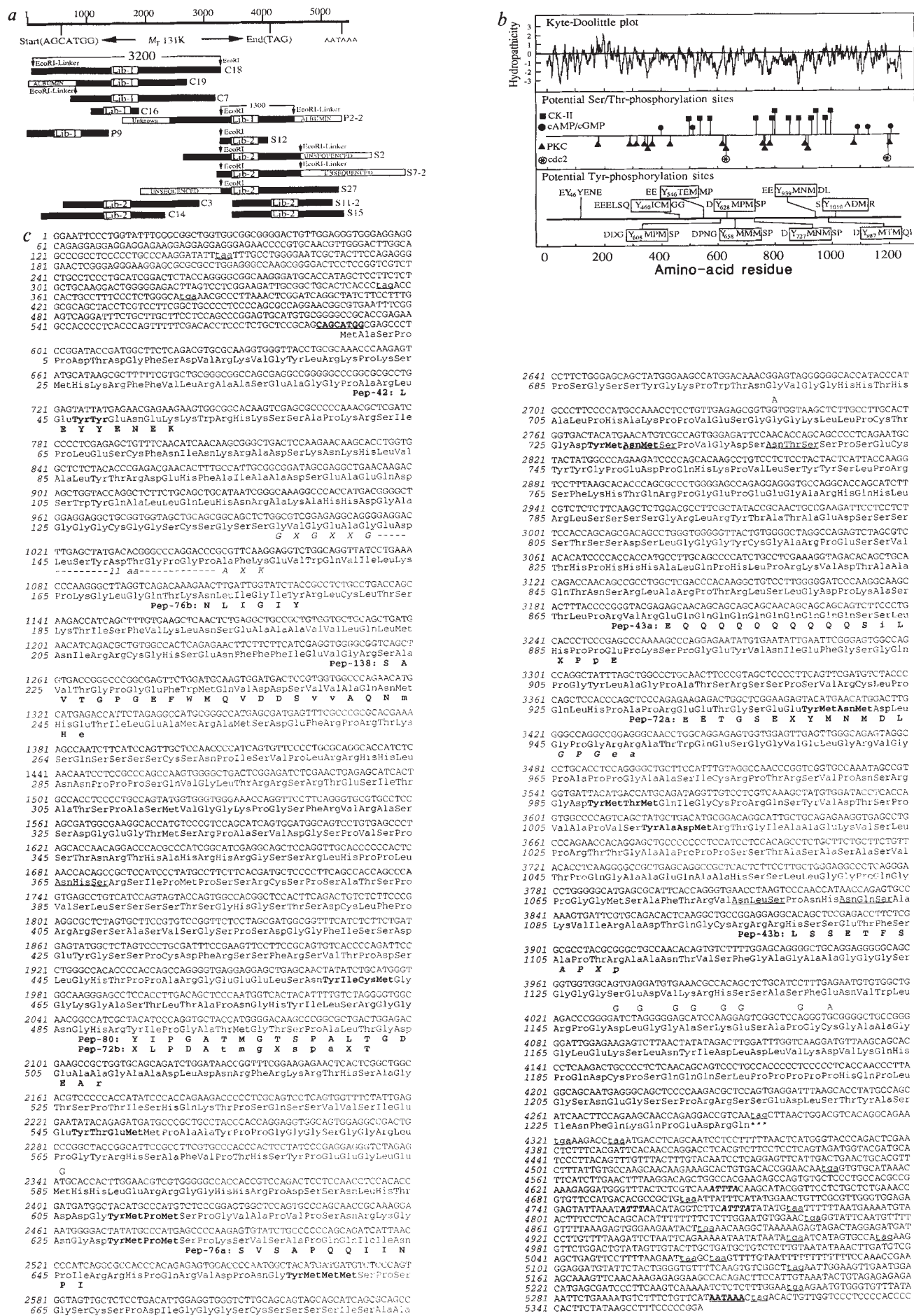
Research Division, Joslin Diabetes Center, Department of Medicine, Harvard Medical School, One Joslin Place, Boston, Massachusetts 02215, USA

SINCE the discovery of insulin nearly 70 years ago, there has been no problem more fundamental to diabetes research than understanding how insulin works at the cellular level. Insulin binds to the α subunit of the insulin receptor which activates the tyrosine kinase in the β subunit, but the molecular events linking the receptor kinase to insulin-sensitive enzymes and transport processes are unknown^{1,2}. Our discovery that insulin stimulates tyrosine phosphorylation of a protein of relative molecular mass between 165,000 and 185,000, collectively called pp185, showed that the insulin receptor kinase has specific cellular substrates³. The pp185 is a minor cytoplasmic phosphoprotein found in most cells and tissues^{4–10}; its phosphorylation is decreased in cells expressing mutant receptors defective in signalling^{6,11}. We have now cloned *IRS-1*, which encodes a component of the pp185 band. *IRS-1* contains over ten potential tyrosine phosphorylation sites, six of which are in Tyr-Met-X-Met motifs. During insulin stimulation, the *IRS-1* protein undergoes tyrosine phosphorylation and binds phosphatidylinositol 3-kinase, suggesting that *IRS-1* acts as a multisite 'docking' protein to bind signal-transducing molecules containing Src-homology 2 and Src-homology-3 domains^{12–14}. Thus *IRS-1* may link the insulin receptor kinase and enzymes regulating cellular growth and metabolism.

We used the partial amino-acid sequence of rat liver pp185 (ref. 15) to prepare optimal complementary DNA probes¹⁶. Two partial cDNA clones (C18 and C19) were initially identified with these probes (Fig. 1a). Further screening yielded 14 overlapping clones that encode *IRS-1*. *IRS-1* is a hydrophilic protein of relative molecular mass 131,000 (M_r , 131K) which contains nine of the original eleven tryptic peptide sequences obtained from rat liver pp185 (Fig. 1b and c). Northern analysis indicates that *IRS-1* messenger RNA is about 9.5 kilobases (data not shown). No Src-homology domains 2 and 3 (SH2/SH3) have been identified in *IRS-1*. But it contains a potential ATP-binding site beginning with a glycine-rich motif (Gly₁₃₇-Val-Gly-Glu-Ala-Gly) and followed by an essential lysine residue 14 amino acids away (Ala-X-Lys₁₅₆-X-Ile, where X is any amino-acid residue) (Fig. 1c)^{17,18}. However, *IRS-1* lacks the Asp-Phe-Gly and Ala-Pro-Glu motifs diagnostic of a protein kinase¹⁷. Moreover, *IRS-1* does not undergo autophosphorylation in immunocomplexes, suggesting that it is not a protein kinase (data not shown).

IRS-1 contains many potential phosphorylation sites. Based on typical motifs for cyclic AMP-dependent protein kinase (R/K-R/K-X-S/T; one-letter amino acid code), protein kinase C (S/T-X-R/K), casein kinase II (S/T-X-X-E/D) and the cdc2 kinase (S/T-P-X-K/R), 35 putative Ser/Thr phosphorylation sites are distributed throughout the protein (Fig. 1b). At least 10 potential tyrosine phosphorylation sites exist, six of which are located in the central region of *IRS-1* and contain the YMXM motif (Fig. 1b); three others have a YXXM motif, and one site has the sequence EYYE. Synthetic peptides containing the YMXM motifs were phosphorylated by the purified

* Present address: Department of Pathology and Laboratory Medicine, University of Pennsylvania School of Medicine, B101 Richards Bldg, Philadelphia, Pennsylvania 19104, USA.



insulin receptor with Michaelis constants, K_m , of about 50 μ M, suggesting that these tyrosine residues are possible phosphorylation sites of IRS-1 *in vivo* (data not shown).

The predicted M_r of IRS-1 (131K) is smaller than expected for a protein that migrates between 175 and 185K during SDS-PAGE. But expression of IRS-1 cDNA in CHO cells

FIG. 1 a, IRS-1 cloning strategy. The cDNA clones C18 and C19 were identified from Lib-1 (see Methods) with a mixture of optimal cDNA probes constructed from the partial amino-acid sequence of pp185 (ref. 15). The remaining clones were identified by screening Lib-1 and Lib-2 with the radiolabelled 3,200-base pair (bp) *Eco*RI insert of clone C18, or the 1,300-bp *Eco*RI fragment from clone P2-2. Overlapping cDNA sequences are indicated by the solid bars, and concatamers formed between cDNAs encoding IRS-1 and albumin or unknown sequences are indicated by the open bars. The *Eco*RI site is found in the cDNA, whereas the *Eco*RI-linkers were contributed by the λ ZapII vector. The partial cDNA is 5,365 bp long and contains an open reading frame which extends from a Kozak start site at nucleotide 589, to the first TAG stop codon at nucleotide 4,293. b, IRS-1 structural features. The hydropathicity of the deduced sequence of IRS-1 was analysed by the Kyte-Doolittle algorithm. The relative location of potential Ser/Thr phosphorylation sites for casein kinase II (CK-II) (S/T-X-X-E/D) cAMP/cGMP kinase (R/K-R/K-X-S/T), protein kinase C (PKC) (S/T-X-R/K), and *cdc2* kinase (S/T-P-X-K/R) are shown. Potential tyrosine phosphorylation sites and the surrounding amino-acid sequence is shown to illustrate the distribution of YMXM or YXXM motifs. c, Complementary DNA and deduced protein sequence of IRS-1. The partial cDNA sequence of IRS-1 is shown with the deduced amino-acid sequence of its 131 K open reading frame. In-frame stop codons in the 5'-untranslated end are indicated in lower-case letters, and the Kozak initiation site, CAGCATGG, is underlined in bold starting at nucleotide 585. The locations of the tryptic peptides obtained from the pp185 band¹⁵ are indicated in bold (single-letter code) under the corresponding amino-acid sequence. The putative ATP-binding site beginning at Gly137 is shown in italic under the corresponding amino-acid sequence. Potential tyrosine phosphorylation sites are indicated by bold typeface in the deduced amino-acid sequence; potential sites of glycosylation are underlined. In the 3'-untranslated region, inframe stop codons are indicated in lowercase, a putative poly(A)⁺ adenylation signal, TATAAA, is underlined in bold, and the mRNA destabilization consensus sequence, ATTAA, is shown in bold italics. Disagreements between the nucleotide sequence of Lib-1 and Lib-2 are indicated by the discordant nucleotide from Lib-2 above the consensus sequence.

METHODS. Two oligo(dT) and random-primed bacteriophage cDNA libraries, Stratagene 936507 (Lib-1) and 936512 (Lib-2) were screened with optimal oligonucleotide probes to Pep 80 and Pep 138. A pair of oligonucleotides (Oligos Etc., Connecticut) with a 12-nucleotide overlap (underlined) were synthesized for Pep 80 (TCTGCTGTGACAGGCCCGCGAGTCTTGGATGCAGGTGG and CATGTTCTGGGCCTTCACAGAGTCATCCACCTGCATCC), and Pep 138 (CCGGGCTCATCGCTGTGAGGCGAGGGAGGTGCCCAT and TACATCCCTGGCGCCACCATGGGCACCTC). Each pair of oligonucleotides (0.6 pmol) was annealed in 10 μ l labelling buffer (Amersham). [³²P]dCTP (210 μ l 1 mCi ml⁻¹, 3,000 Ci per mmol) and [³²P]dGTP (21 μ l 20 mCi ml⁻¹, 6,000 Ci per mmol) were mixed and lyophilized in microfuge tubes, followed by addition of 26 μ l H₂O, 4 μ l of 5 $\times$ labelling buffer, 4 μ l dATP and dTTP, 10 μ l annealed oligos, and extended with excess Klenow (Amersham). The mixture was incubated at room temperature for 2 h and then at 37 °C for 30 min. The labelled probes were separated from free dNTPs using an Elutip (Schleicher & Schuell). Specific activity was 2 $\times$ 10⁸ c.p.m. per pmol for probe 80 and 2.25 $\times$ 10⁹ c.p.m. per pmol for probe 138. About 1.5 $\times$ 10⁶ plaques were plated at a density of 50,000 plaques per 150-mm plate, transferred to nitrocellulose filters (Schleicher & Schuell), and screened with an equimolar mixture of probes 80 and 138 (3 $\times$ 10⁶ c.p.m. per ml). Hybridizations were performed overnight in 5 $\times$ Denhardt's solution containing 20% formamide, 10% dextran sulphate, 6 $\times$ SSC, and 50 mM sodium phosphate (pH 6.8) containing 100 μ g ml⁻¹ salmon sperm DNA. Filters were washed 3 times with 2 $\times$ SSC containing 0.1% SDS at 22 °C for 30 min, then with 0.2 $\times$ SSC and 0.1% SDS for 30 min at 37 °C. Dried blots were exposed to Kodak XAR-5 film with a Quanta 111 intensifying screen at -70 °C. The pBluescript SK⁻ plasmid containing the cDNA inserts that remained positive after two rounds of plaque purification were released from the λ ZapII vector by *in vivo* excision with the helper phage R408, as described in the manufacturers instructions (Stratagene). Inserts were sequenced on both strands with Sequenase (USB) using specific primers selected at convenient intervals. The sequence was confirmed by sequencing the coding strand of independent cDNA inserts obtained from Lib-2. Sequences were aligned and analysed using the EUGENE and SAM programs (Molecular Biology Computing Research Resource, Dana Faber Cancer Institute and Harvard School of Public Health).

(CHO/IRS-1) revealed that IRS-1 migrates as a 180K phosphoprotein during insulin stimulation. IRS-1 was immunoprecipitated from insulin-stimulated CHO/IRS-1 cells with anti-phosphotyrosine antibody, and it migrated exactly with pp185 from CHO/IR cells (Fig. 2, lanes a-d). The pp185 was undetected in CHO/neo cells during insulin stimulation (Fig. 2, lanes e and f), suggesting that either overexpression of IRS-1 or the insulin receptor was necessary to obtain a detectable signal in this assay.

The juxtamembrane region of the insulin receptor β subunit is essential for signal transmission, and mutations in this region by substitution of Tyr 960 with phenylalanine (CHO/IR_{F960}), or deletion of 12 amino acids around Tyr 960 (CHO/IR _{Δ 960}), impairs normal insulin-stimulated tyrosine phosphorylation of pp185 (refs 19, 20). As shown previously, the mutant receptors autophosphorylate normally during immunoprecipitation with the anti-phosphotyrosine antibody, but pp185 was immunoprecipitated only from the insulin-stimulated CHO/IR cells (Fig. 3a, lanes a and b). Using anti-IRS-1 antibody, a single [³²P]phosphoprotein was identified in each cell line before insulin stimulation which migrated at a position corresponding to an M_r of 165-175K (Fig. 3b). This protein was not immunoprecipitated with anti-phosphotyrosine antibody as it

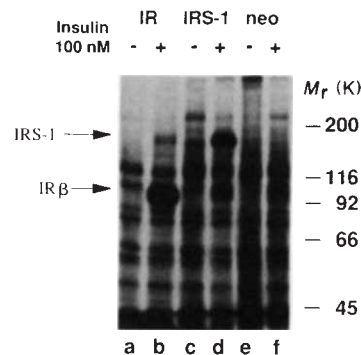
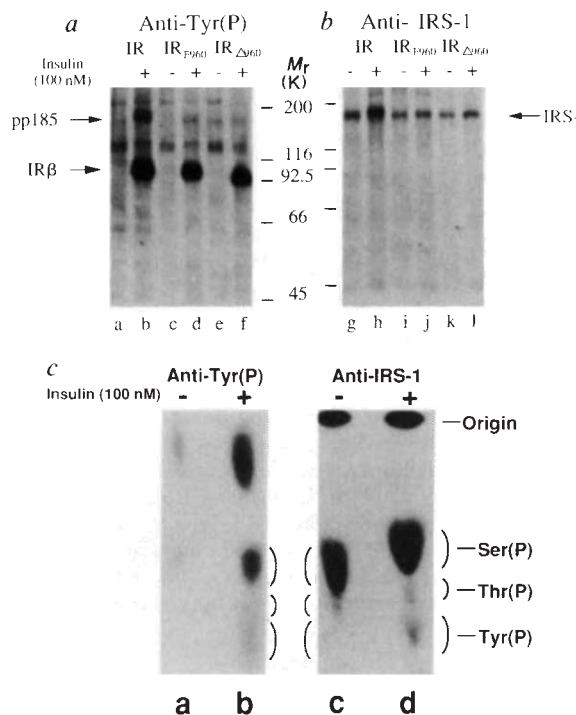


FIG. 2 Expression and tyrosine phosphorylation of IRS-1 in CHO cells. CHO cells expressing 10⁶ insulin receptors (CHO/IR) (lanes a and b), IRS-1 (CHO/IRS-1) (lanes c and d), or no exogenous cDNA (CHO/neo) (lanes e and f) were labelled with [³²P]orthophosphate¹⁹ and incubated without (a, c and e) or with 100 nM insulin for 1 min (b, d and f). Phosphotyrosine-containing proteins were immunoprecipitated from cell extracts with antiphosphotyrosine antibody separated by SDS-PAGE under reducing conditions, and detected by autoradiography¹⁹. The migration positions of the β subunit of the insulin receptor and IRS-1 are indicated.

METHODS. The CHO-IR and CHO/neo cells have been described¹⁹. CHO/IRS-1 cells were prepared by calcium phosphate-mediated transfection of the IRS-1 cDNA in the expression vector pCMVhis (M. Birnbaum, Harvard Medical School) which contains a histidinol resistance gene²⁸. The noncoding region of the IRS-1 cDNA contains several inframe start and stop codons which might interfere with the efficient translation. These regions were removed as follows: the sequence of 553 to 997, including the start codon at 588 and *Bst* EII sit at 642 was amplified by polymerase chain reaction (PCR). A 5'-end primer located at position 553 adapted with an *Spe*I site (TCAACTAGTTTTTCGACACCTCCCTCTGCT) and 3'-end primer at nucleotide 997 (CAGAGCTGCCGCTGCA) in the IRS-1 cDNA were synthesized (Oligos Etc.). PCR was performed in 100 μ l 100 mM Tris-HCl, pH 8.3, containing 40 mM KCl, 1.5 mM MgCl₂, 0.01% gelatin, 0.2 mM dNTP, 50 pmol of primers, 0.5 μ g full-length IRS-1 cDNA in pBluescript-II and 2.5 units of *Taq* DNA polymerase (Perkin Elmer Cetus). The reaction was cycled 10 times at 94 °C (1 min), 55 °C (2 min) and 72 °C (1.5 min). Both PCR products and full-length pp185 cDNA in pBluescript (Stratagene) were cut with *Spe*I and *Bst*EII. The 5'-end region in the full-length IRS-1 cDNA was replaced by the fragment of 553-642 released from the PCR product. The new vector carrying the modified IRS-1 cDNA was confirmed by sequencing and restriction mapping. The complete IRS-1 cDNA was confirmed by DNA sequencing and released from pBluescript by digestion with *Spe*I and *Eco*RV, blunt-end inserted into pCMVhis at a blunt-ended *Hind*III site. CHO cells were co-transfected with this plasmid (10 μ g) and pSVneo (0.5 μ g) selected with 800 μ g ml⁻¹ G418 (ref. 19), and then with 10 mM histidinol (Sigma). Four surviving clones were selected and all expressed raised levels of IRS-1.



contained predominantly phosphoserine and a small amount of phosphothreonine (Fig. 3c). After insulin stimulation of the CHO/IR cells, phosphorylation of IRS-1 doubled and it migrated with a higher M_r (175–185K) (Fig. 3b, lanes g and h). IRS-1 and pp185 from insulin-stimulated CHO/IR cells contained phosphotyrosine (Fig. 3c). Insulin had little or no effect on the phosphorylation of IRS-1 in the CHO/IR_{F960} and CHO/IR_{Δ960} cells, which is consistent with the absence of pp185 in the corresponding anti-phosphotyrosine immunoprecipitates (Fig. 3a and b). These results show that IRS-1 has characteristics similar to those of pp185, suggesting that they are related proteins.

A phosphatidylinositol 3'-kinase (PtdIns 3'-kinase) associates with certain phosphotyrosine-containing proteins, including receptors for platelet-derived growth factor (PDGF), epidermal growth factor, colony stimulating factor-1 and insulin^{12,21–23}. These ligands also increase the amount of PtdIns 3-phosphate in intact cells¹² which may be important in growth control¹². As previously described^{22,24}, PtdIns 3'-kinase was detected in anti-phosphotyrosine and anti-insulin receptor immunoprecipitates

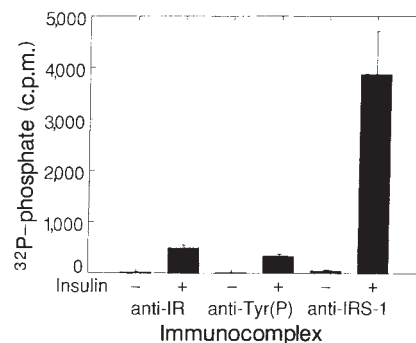
FIG. 3 Insulin-stimulated phosphorylation of pp185 and IRS-1 in CHO/IR cells, or CHO cells expressing mutant receptors (IR_{F960} or IR_{Δ960}) has been previously described^{19,20}. CHO/IR cells (lanes a, b, g and h), CHO/IR_{F960} cells (lanes c, d, i and j) and CHO/IR_{Δ960} cells (lanes e, f, k and l) were labelled with 0.5 mCi ml⁻¹ [³²P]orthophosphate for 2 h, and then incubated without (–) insulin or with (+) 100 nM insulin for 1 min¹⁹. Cell extracts were prepared and immunoprecipitated with anti-phosphotyrosine antibody (a) or the specific IRS-1 antibody (b), separated by SDS-PAGE under reducing conditions, and the phosphoproteins detected by autoradiography. c, Phosphoamino-acid composition of pp185 and IRS-1 immunoprecipitated from CHO/IR cells with anti-phosphotyrosine antibody (lanes a and b) or anti-IRS-1 (lanes c and d), respectively. Phosphoamino-acid analysis was carried out on tryptic digests²⁹. The anti-IRS-1 antibody was prepared in rabbits with synthetic Pep80 (provided by S. Shoelson, Joslin Diabetes Center) coupled to keyhole limpet haemocyanin³⁰.

after stimulation of CHO/IR cells, but eight times more PtdIns 3'-kinase activity was found in anti-IRS-1 immunoprecipitates (Fig. 4). Immunoprecipitation of PtdIns 3'-kinase by anti-IRS-1 antibody was blocked by the peptide antigen and did not occur with nonspecific rabbit antibodies; HPLC analysis confirmed that the principal product was phosphatidylinositol 3-phosphate (data not shown).

The association of PtdIns 3'-kinase with IRS-1 during insulin stimulation supports the hypothesis that IRS-1 binds signal transduction molecules during insulin-stimulated tyrosine phosphorylation. It is unlikely that IRS-1 itself is a PtdIns 3'-kinase, because the purified PtdIns 3'-kinase exists as a 110K catalytic subunit and an 85K subunit which contains two SH2 domains and one SH3 domain^{13,21,25,26}. The association between PtdIns 3'-kinase and IRS-1 probably occurs through phosphorylated YMXM motifs on IRS-1 and the SH2/SH3 domains on the 85K subunit. Insulin increases the cellular concentration of PtdIns 3-phosphate, suggesting that the PtdIns 3'-kinase is activated during insulin stimulation²². Tyrosine phosphorylation of the 85K subunit may activate the kinase²⁷, but there is no

FIG. 4 Insulin stimulation of phosphatidylinositol 3'-kinase. CHO/IR cells were stimulated with 100 nM insulin for 10 min and extracted. The PtdIns 3'-kinase activity was assayed in immunocomplexes prepared with antiphosphotyrosine antibody (anti-Tyr(P)), anti-insulin receptor antibody (anti-IR)³¹, and anti-IRS-1 antibody.

METHODS. *In vitro* phosphorylation of phosphatidylinositol was carried out in the immunocomplexes as described²². Subconfluent CHO cells grown in 100-mm dishes were made quiescent by an overnight incubation in F-12 medium containing 0.5% BSA. The cells were then incubated in the absence or presence of insulin (100 nM) for 10 min, and washed once with ice-cold PBS and twice with 20 mM Tris (pH 7.5) containing 137 mM NaCl, 1 mM MgCl₂, 1 mM CaCl₂, and 100 μM Na₃VO₄ (buffer A). The cells were solubilized in 1 ml buffer A containing 1% N-P40 (Sigma) and 10% glycerol, and insoluble material was removed by centrifugation at 13,000g for 10 min. Immunocomplexes were precipitated from the supernatant with protein A-Sepharose (Pharmacia) and washed successively in (1) PBS containing 1% N-P40 and 100 μM Na₃VO₄ (3 times); (2) 100 mM Tris (pH 7.5) containing 500 mM LiCl₂ and 100 μM Na₃VO₄ (3 times); and (3) 10 mM Tris (pH 7.5) containing 100 mM NaCl, 1 mM EDTA and 100 μM Na₃VO₄ (twice). The pellets were resuspended in 50 μl 10 mM Tris (pH 7.5) containing 100 mM NaCl and 1 mM EDTA. To each pellet was added 10 μl 100 mM MgCl₂ and 10 μl phosphatidylinositol (Avanti) (2 μg μl⁻¹) sonicated in 10 mM Tris (pH 7.5), 1 mM EGTA. The reaction was started by the addition of 10 μl 440 μM ATP containing 30 μCi



[³²P]ATP. After 10 min at 22 °C, the reaction was stopped by the addition of 20 μl 8 M HCl and 160 μl CHCl₃:methanol (1:1). Samples were centrifuged, and the lower organic phase removed and applied to a silica gel thin-layer chromatography plate (Merck) which had been coated with 1% potassium oxalate. Thin-layer chromatography plates were developed in CHCl₃:CH₃OH:H₂O:NH₄OH (60:47:11:3:2), dried and visualized by autoradiography. The radioactivity in spots which co-migrated with PtdIns-4P standard (Sigma) was measured by Cerenkov counting as before²².

evidence for phosphorylation of an 85K band during insulin stimulation. Activation may occur during the binding of the PtdIns 3'-kinase to IRS-1, but there is no direct evidence for this. The PtdIns 3'-kinase binds weakly to the insulin receptor, which may be an important step to recruit the enzyme to the plasma membrane. Recovery of the PtdIns 3'-kinase in anti-insulin receptor immunoprecipitates, however, could be due to an association between the insulin receptor and the IRS-1/PtdIns 3'-kinase complex, rather than direct binding of the insulin receptor to the PtdIns 3'-kinase.

Other signal transduction proteins, including the phosphoinositide-specific phospholipase $C_{\gamma 1}$, the GTPase-activating protein and various Src-like tyrosine kinases, contain SH2/SH3 domains and associate strongly with certain membrane-bound phosphotyrosine-containing proteins, such as the PDGF receptor¹²⁻¹⁴. The PDGF receptor contains a YMXM and a homologous YVXM motif in its kinase-insert region which is essential for binding PtdIns 3'-kinase¹². The association between the PDGF receptor and the PtdIns 3'-kinase, but not phospholipase $C_{\gamma 1}$ or GTPase-activating protein, is blocked with phosphopeptides containing the YMXM motif, suggesting that phosphorylated YMXM motifs form specific recognition sites for proteins containing certain isoforms of the SH2/SH3 domain^{12,21}. The presence of nine potential tyrosine phosphorylation sites in YMXM and YVXM motifs suggests that IRS-1 may act as a multisite 'docking' protein which binds a variety of signal-transducing molecules that contain the appropriate SH2/SH3 domains. Other tyrosine kinases may also phosphorylate specific sites in IRS-1, and serine/threonine kinases could regulate association between IRS-1 and other molecules. This model may begin to explain the pleiotropic effects of insulin, especially if the activity or cellular location of various signal transduction molecules is altered during association with tyrosine-phosphorylated IRS-1. We have been unable to identify an intrinsic enzymatic activity for IRS-1, but if one exists it could also play an important part in insulin signalling. □

Androgen receptor gene mutations in X-linked spinal and bulbar muscular atrophy

Albert R. La Spada*, Elizabeth M. Wilson†, Dennis B. Lubahn†, A. E. Harding‡ & Kenneth H. Fischbeck*

* Neurology Department, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania 19104-6146, USA

† Laboratories for Reproductive Biology, University of North Carolina, Chapel Hill, North Carolina 27599, USA

‡ Institute of Neurology, The National Hospital, Queen Square, London WC1N 3BG, UK

X-LINKED spinal and bulbar muscular atrophy (Kennedy's disease) is an adult-onset form of motorneuron disease which may be associated with signs of androgen insensitivity. We have now investigated whether the androgen receptor gene on the proximal long arm of the X chromosome is a candidate gene for this disease. In patient samples we found androgen receptor gene mutations with increased size of a polymorphic tandem CAG repeat in the coding region. These amplified repeats were absolutely associated with the disease, being present in 35 unrelated patients and none of 75 controls. They segregated with the disease in 15 families, with no recombination in 61 meioses (the maximum log likelihood ratio (lod score) is 13.2 at a recombination rate of 0). The association is unlikely to be due to linkage disequilibrium, because 11 different disease alleles were observed. We conclude that enlargement of the CAG repeat in the androgen receptor gene is probably the cause of this disorder.

X-linked spinal and bulbar muscular atrophy (SBMA) is characterized by progressive muscle weakness and atrophy¹, but affected males also may have gynecomastia and reduced fertility, suggesting a defect in androgen receptor function². A recent study showed decreased androgen binding to cultured fibroblasts in one family³. The human androgen receptor (AR) gene

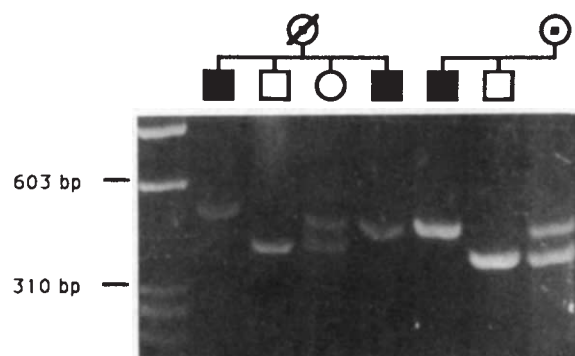


FIG. 1 Agarose-gel electrophoresis of PCR products obtained by amplification of the AR CAG repeat for two unrelated SBMA pedigrees. Affected individuals have enlarged fragments (~500 base pairs (bp)), whereas their unaffected brothers have fragment sizes within the expected normal range (405–450 bp). Obligate heterozygous (carrier) females have both an enlarged fragment and a normal fragment.

METHODS. PCR was performed in a final volume of 50 μ l containing 500 ng genomic DNA, 25 pmol of each oligonucleotide primer (5'-GCCTGTTGAACTCTTCTGAGC-3' and 5'-GCTGTGAAGGTTGCTGTCTCTC-3'), 200 μ M each dNTP, 50 mM KCl, 10 mM Tris-HCl (pH 8.3), 1.5 mM MgCl₂, 0.01% gelatin and 4 U AmpliTaq (Perkin-Elmer Cetus). The reaction mixtures were denatured at 95 °C for 5 min before the addition of Taq polymerase. Thirty-five temperature cycles (each for 1 min at 95 °C, 2 min at 65 °C and 1.5 min at 72 °C) were carried out in a Coy Tempcycler. The final extension step lasted 8.5 min. One-quarter of the amplified DNA for each sample was analysed by electrophoresis on a 1.2% agarose gel containing ethidium bromide with ϕ X174HaeIII-digested DNA as a size standard.

Received 18 January; accepted 9 May 1991.

- Kahn, C. R. & White, M. F. *J. clin. Invest.* **82**, 1151–1156 (1988).
- Rosen, O. M. *Science* **237**, 1452–1458 (1987).
- White, M. F., Maron, R. & Kahn, C. R. *Nature* **318**, 183–186 (1985).
- Gibbs, E. M., Allard, W. J. & Leinhard, G. E. *J. biol. Chem.* **261**, 16597–16603 (1986).
- Maegawa, H. *et al. J. biol. Chem.* **263**, 12629–12637 (1988).
- White, M. F., Stegmann, E. W., Dull, Y. J., Ullrich, A. & Kahn, C. R. *J. biol. Chem.* **262**, 9769–9777 (1987).
- Kadowaki, T. *et al. J. biol. Chem.* **262**, 7342–7350 (1987).
- Shemer, J. *et al. J. biol. Chem.* **262**, 15476–15482 (1987).
- Momomura, K., Tobe, K., Seyama, Y., Takaku, F. & Kasuga, M. *Biochem. biophys. Res. Commun.* **155**, 1181–1186 (1988).
- Tobe, K. *et al. Diabetes* **39**, 528–533 (1990).
- Wilden, P. A. *et al. Proc. natn. Acad. Sci. U.S.A.* **87**, 3358–3362 (1990).
- Cantley, L. C. *et al. Cell* **64**, 281–302 (1991).
- Escobedo, J. A. *et al. Cell* **65**, (1991).
- Anderson, D. *et al. Nature* **250**, 979–982 (1990).
- Rothenberg, P. L. *et al. J. biol. Chem.* **266**, 8302–8311 (1991).
- Lathe, R. *J. molec. Biol.* **83**, 1 (1985).
- Hanks, S. K., Quinn, A. M. & Hunter, T. *Science* **241**, 42–52 (1990).
- Pines, J. & Hunter, T. *The New Biologist* **2**, 389–401 (1990).
- White, M. F. *et al. Cell* **54**, 641–649 (1988).
- Backer, J. M. *et al. Biochemistry* (in the press).
- Escobedo, J. A., Kaplan, D. R., Kavanaugh, W. M., Turck, C. W. & Williams, L. T. *Molec. cell. Biol.* **11**, 1125–1132 (1991).
- Ruderman, N., Kapeller, R., White, M. F. & Cantley, L. C. *Proc. natn. Acad. Sci. U.S.A.* **87**, 1411–1415 (1990).
- Endemann, G., Yonezawa, K. & Roth, R. A. *J. biol. Chem.* **265**, 396–400 (1990).
- Hjollund, E., Richelsen, B. & Pedersen, O. *Acta endocrinol. (Copenh)* **118**, 59–67 (1988).
- Morgan, S. J., Smith, A. D. & Parker, P. J. *Eur. J. Biochem.* **191**, 761–767 (1990).
- Carpenter, C. L. *et al. J. biol. Chem.* **265**, 19704–19711 (1990).
- Kaplan, D. R. *et al. Cell* **50**, 1021–1029 (1987).
- Hartman, S. C. & Mulligan, R. C. *Proc. natn. Acad. Sci. U.S.A.* **85**, 8047–8051 (1988).
- Haring, H.-U., Kasuga, M., White, M. F., Crettaz, M. & Kahn, C. R. *Biochemistry* **23**, 3298–3306 (1984).
- Perlman, R., Bottaro, D., White, M. F. & Kahn, C. R. *J. biol. Chem.* **264**, 8946–8950 (1989).
- Brindle, N. P. J., Tavare, J. M., Dickens, M., Whittaker, J. & Siddle, K. *Biochem. J.* **268**, 615–620 (1990).

ACKNOWLEDGEMENTS. We thank S. A. Kahn, M. G. Myers Jr, and M. Miralpeix for their help throughout this project. This work was supported in part by the NIH (C.R.K. and M.F.W.), the Juvenile Diabetes Foundation (M.F.W.), and the Diabetes and Endocrinology Research Center, P.A.W. and J.M.B. are of National Research service awards. C.R.K. is a recipient of a Pfizer Biomedical Research Award and M.F.W. is a scholar of the Pew Foundation, Philadelphia.

METHODS. Reaction mixtures containing 250 ng genomic DNA, 12.5 pmol each oligonucleotide primer (5'-TCCAGAATCTGTTCCAGACGGTGC-3' and 5'-GCTGTGAAGGTTGCTGTCTCAT-3') and 4 U AmpliTaq were prepared in the recommended buffer and at appropriate nucleotide concentrations, except that the total volume was decreased to 25 μ l and 2.5 μ Ci of [α - 32 P]dATP was included in each reaction tube. 25 cycles of 95 $^{\circ}$ C for 1 min, 60 $^{\circ}$ C for 1 min and 72 $^{\circ}$ C for 1 min were performed with the initial denaturation step and final elongation step lengthened to 5 min and 8 min, respectively. After addition of 15 μ l stop solution (95% formamide, 20 mM EDTA, 0.05% bromophenol blue, 0.05% xylene cyanol FF), samples were denatured at 75 $^{\circ}$ C for 3 min and then electrophoresed in 4% polyacrylamide/8 M urea gels for 3 h at 38 W. Gels were exposed to Kodak X-Omat film overnight at -70 $^{\circ}$ C.

This repeat falls within the coding domain of the gene, corresponding to a long tract of glutamine residues beginning at position 58. PCR amplification of the DNA segment containing the CAG repeat in samples from SBMA patients gave a band representing fragments that are about 75 base pairs larger than those for unaffected individuals in the same families and controls (Fig. 1). The larger band segregated with the disease in 15 SBMA families, with no recombinations in 61 meioses, and linkage analysis using the computer program LIPED¹⁰ demonstrated a lod score of 13.2 at 0 centimorgans. There was absolute association of the larger-fragment band with the disease phenotype and the smaller-fragment band with normals.



Our results are not likely to be due to linkage disequilibrium between the disease locus and the AR polymorphism, because the SBMA patients had at least 11 different enlarged (CAG) domains which were not seen in more than 300 controls reported here and elsewhere. Furthermore, our 35 SBMA probands comprise an extremely diverse ethnic population, with individuals of English, Irish, Polish, French, Italian, German, Norwegian, Japanese, Chinese, Turkish and Mexican ancestry. Also, a low-frequency *Hind*III restriction fragment length polymorphism (RFLP)⁴, detected with the AR cDNA as probe, enabled us to analyse the genetic background of the SBMA patients. If there were linkage disequilibrium between SBMA and the AR CAG polymorphism, then we would expect association with one allele of the AR RFLP as well; but we found that SBMA patients have both alleles of the *Hind*III RFLP in about the same ratio as controls.

Downloaded from <http://ajphaphysocpharm.sagepub.com/> at 11:06 11 September 2014

METHODS. Dideoxy nucleotide sequencing of the CAG repeat was performed with PCR fragments cut out of agarose gels and isolated by the GeneClean

Enlargement of the CAG repeat thus appears to be specific to X-linked SBMA. We propose that the corresponding increase in the polyglutamine tract of the androgen receptor is responsible for the SBMA phenotype. The androgen receptor has domains for hormone and DNA binding, as do other members of the steroid hormone receptor family. The CAG repeat in the first exon of the AR gene encodes another portion of the receptor protein which is not involved in hormone or DNA binding (Fig. 3). A similar CAG repeat is present in the first exon of the rat androgen receptor gene¹¹. The function of these CAG repeats is unknown, although repeats with this structure, designated *opa* repeats, have been identified at other developmentally regulated loci in *Drosophila* and mouse^{12,13}. Glutamine-rich domains are important for the function of the transcriptional factors Sp1 and Antennapedia¹⁴, and a long polyglutamine tract is present in transcription factor IID¹⁵. The polyglutamine tract encoded by the CAG repeat in the AR gene may be involved in transcriptional regulation. If so, we would predict that the regulation by

AR is altered in SBMA patients, perhaps in a tissue-specific manner.

Absence of androgen receptors leads to testicular feminization, with abnormal fetal sexual differentiation and no weakness; this is distinctly different from SBMA, where differentiation of the genitalia in the fetus is normal, despite later signs of insensitivity to androgens. The abnormality we found in the first exon of the AR gene in SBMA patients may not change DNA or androgen binding enough to produce testicular feminization, although it probably alters AR function in motoneurons. Androgen receptors are normally concentrated in spinal and bulbar motoneurons, the cells that degenerate in SBMA¹⁶. Also, androgens can affect spinal motoneuron growth, development and response to injury in rats^{17,18}. In SBMA, enlargement of the polyglutamine repeat may prevent the androgen receptor from performing an important regulatory activity in motoneurons, thereby leading to the degeneration of these cells which is characteristic of the disease. □

Received 5 March; accepted 28 May 1991.

- Harding, A. E. et al. *J. Neurol. Neurosurg. Psychiatry* **45**, 1012–1019 (1982).
- Arbizu, T. et al. *J. Neurol. Sci.* **59**, 371–382 (1983).
- Warner, C. L. et al. *Neurology* **41** (Suppl. 1), 313 (1991).
- Brown, C. J. et al. *Am. J. hum. Genet.* **44**, 264–269 (1989).
- Fischbeck, K. H. et al. *Neurology* **36**, 1595–1598 (1986).
- Fischbeck, K. H., Souders, D. & La Spada, A. *Adv. Neurol.* **56**, 209–213 (1991).
- Edwards, A., Hammond, H. A., Caskey, C. T. & Chakraborty, R. *Genomics* (submitted).
- Lubahn, D. B. et al. *Molec. Endocrinol.* **2**, 1265–1275 (1988).
- Tilley, W. D. et al. *Proc. natn. Acad. Sci. U.S.A.* **86**, 327–331 (1989).
- Ott, J. *Am. J. hum. Genet.* **26**, 588–597 (1974).
- Tan, J. et al. *Molec. Endocrinol.* **2**, 1276–1285 (1988).
- Wharton, K. A., Yedvobnick, B., Finnerty, V. G. & Artavanis-Tsakonas, S. *Cell* **40**, 55–62 (1985).

- Duboule, D., Haenlin, M., Galliot, B. & Mohier, E. *Molec. cell. Biol.* **7**, 2003–2006 (1987).
- Mitchell, P. J. & Tjian, R. *Science* **245**, 371–378 (1989).
- Kao, C. C. et al. *Science* **248**, 1646–1649 (1990).
- Sar, M. & Stumpf, W. E. *Science* **197**, 77–80 (1977).
- Kurz, E. M., Sengelaub, D. R. & Arnold, A. P. *Science* **232**, 395–398 (1986).
- Yu, W. A. *J. Neurosci.* **9**, 3908–3914 (1989).
- Yandell, D. W. & Dryja, T. P. *Am. J. hum. Genet.* **45**, 547–555 (1989).

ACKNOWLEDGEMENTS. We thank the families who participated in this study and the following clinicians who contributed patient samples: I. Hausmanowa-Petrusewicz, S. Ringel, L. Stern, C. Warner, J. Mikol, J. Pouget, G. Serratrice, A. Danek, K. Zerres, L. Venier, D. Jacoby, G. Nicholson, N. Laing, T. Munsat, W. Bradley and A. Boughey. This work was supported by the Muscular Dystrophy Association, the March of Dimes Birth Defects Foundation, and the NIH. D.B.L. is a Pew Scholar.

Structure of recombinant human rheumatoid arthritic synovial fluid phospholipase A₂ at 2.2 Å resolution

J.-P. Wery, R. W. Schevitz, D. K. Clawson, J. L. Bobbitt, E. R. Dow, G. Gamboa, T. Goodson, Jr., R. B. Hermann, R. M. Kramer, D. B. McClure, E. D. Mihelich, J. E. Putnam, J. D. Sharp, D. H. Stark, C. Teater, M. W. Warrick & N. D. Jones

Lilly Research Laboratories, Lilly Corporate Center, Eli Lilly and Company, Indianapolis, Indiana 46285, USA

PHOSPHOLIPASES A₂ (PLA₂s) may be grouped into distinct families of proteins that catalyse the hydrolysis of the 2-acyl bond of phospholipids and perform a variety of biological functions. The best characterized are the small (relative molecular mass ~14,000) calcium-dependent, secretory enzymes of diverse origin, such as pancreatic and venom PLA₂s¹. The structures and functions of several PLA₂s are known². Recently, high-resolution crystal structures of complexes of secretory PLA₂s with phosphonate phospholipid analogues have provided information about the detailed stereochemistry of transition-state binding³, confirming the proposed catalytic mechanism of esterolysis⁴. By contrast, studies on mammalian nonpancreatic secretory PLA₂s (s-PLA₂s) have only recently begun; s-PLA₂s are scarce in normal cells and tissues but large amounts are found in association with local and systemic inflammatory processes and tissue injury in animals and man^{5–7}. Such s-PLAs have been purified from rabbit and rat inflammatory exudate^{8,9}, from synovial fluid from patients with rheumatoid arthritis^{10,11} and from human platelets¹¹. Cloning and sequencing shows that the primary structure of the human s-PLA₂^{11,12} has about 37% homology with that of bovine pancreatic

PLA₂¹³ and 44% homology with that of *Crotalus atrox* PLA₂¹⁴. The human s-PLA₂ is an unusually basic protein, yet contains most of the highly conserved amino-acid residues and sequences characteristic of the PLA₂s sequenced so far^{1,15}. Here we report the refined, three-dimensional crystal structure at 2.2 Å resolution of recombinant human rheumatoid arthritic synovial fluid PLA₂. This may aid the development of potent and specific inhibitors of this enzyme using structure-based design.

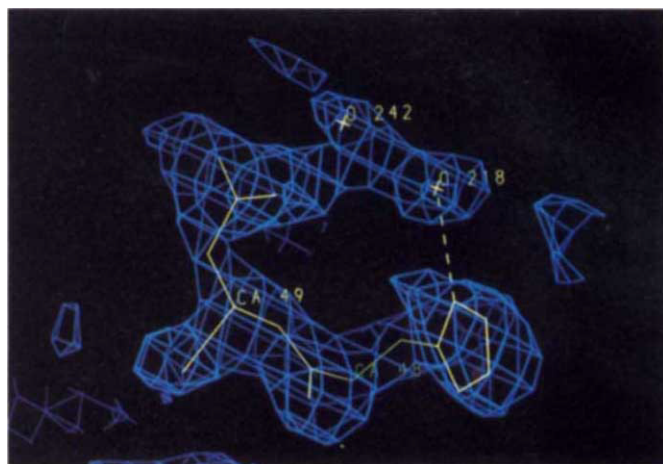
The details of the cloning, expression, isolation, purification, crystallization and structure determination are given in the legend to Fig. 1. A comparison of the primary structure of human s-PLA₂ with the entire family of PLA₂s has previously shown it to be most like group II PLA₂s¹¹. It has the same characteristic disulphide pattern, a small loop at residues 57–68 (numbered as in ref. 11) and a 7-residue C-terminal extension. When the α -carbon positions of human s-PLA₂ are superimposed on those of bovine pancreatic¹⁶ (group I) and *Crotalus atrox*¹⁷ (group II) PLA₂, root-mean-square differences of 1.93 and 1.63 Å, respectively, are observed, further indicating the greater similarity to group II PLA₂s. The greatest differences occur between the human s-PLA₂ and the bovine pancreatic enzyme in loop 57–68, a region of high structural diversity and where the latter has an additional six residues forming a short helix (Fig. 2).

Although human s-PLA₂ resembles the *C. atrox* enzyme more than bovine pancreatic PLA₂, it still shows notable differences. Loop 57–68 is two residues larger and more protruding in the human enzyme. In addition, the seven residue C-terminal extension shows large positional differences along its length, exceeding 5 Å at Gly 127. The large β arm (78–89) also differs from *C. atrox*; however, in this case it is more like the bovine pancreatic PLA₂, except for residues 80 and 81, where the human enzyme differs from both.

Our crystals of human s-PLA₂ do not contain calcium. This may be because the crystals were grown at pH 5.5 in citrate buffer. But five of the seven expected calcium ion ligands are still poised in roughly the position needed for binding. Four of

FIG. 1 Representative portion of a $(2F_o - F_c)$ electron density map of human s-PLA₂ showing a small part of the active site region, His 48, Asp 49 and the oxygen atoms of water molecules Wat 218 and Wat 242.

METHODS. Human s-PLA₂ was produced in Syrian hamster AV12 cells using the adenovirus-based DNA vector pH₂₂ driving expression of two ligated, polymerase chain reaction-amplified human genomic DNA fragments (our unpublished results) based on the published sequence¹¹. Transfected clones were selected and amplified using methotrexate (up to 32 μ M), and assayed for secretion of enzymatically active PLA₂ (ref. 11). One of the best (clone A28) of 96 such clones was grown in an Opticell bioreactor (model 5300, Charles River Biotechnology), which was outfitted with an S-1251 core and operated in a batch mode for large-scale production of s-PLA₂-containing 'conditioned medium' (output of 15–20 l day⁻¹). Both scale-up and production phases were done using a modified mixture of Dulbecco's Modified Eagle's Medium (DME) and Ham's F-12 Medium supplemented with 5% (v/v) defined fetal bovine serum (Hyclone FBS). More than 2,100 litres of medium containing enzyme at a concentration of 5 to 250 ng ml⁻¹ was collected and processed in batches of 50 to 100 litres. The PLA₂ was isolated by a combination of cation exchange chromatography, gel filtration and high pressure reversed-phase chromatography¹¹ with an overall yield of >50%. The final freeze-dried product was >99% pure on rechromatography by analytical HPLC and was homogeneous upon gradient SDS gel electrophoresis (loading >25 μ g, Coomassie blue stain). The enzyme was assayed using autoclaved [³H] oleic acid-labelled *Escherichia coli* as substrate¹¹. A single crystal (150 $\times$ 250 $\times$ 800 μ m) of human s-PLA₂ was grown by vapour diffusion in about 2 weeks from a solution containing 10.0 mg ml⁻¹ protein in 100 mM citrate, pH 5.5, and ammonium chloride at a concentration of 4.0 M. The space group is $P2_1$ with one molecule in the asymmetric unit. The unit cell dimensions are: $a = 31.6$, $b = 54.6$, $c = 31.3$ Å and $\beta = 97.2^\circ$. A full diffraction data set was collected on a Xentronics area detector to 2.2 Å resolution with copper K α radiation. The intensity data was processed using the program XENGEN²³, with a resulting $R_{\text{sym}} = 6.3\%$. To solve the structure of the human enzyme by the molecular replacement technique, a starting model was constructed by 'humanizing' the *C. atrox* structure¹⁷. The human s-PLA₂ is 44% homologous with the *C. atrox* enzyme and has two additional residues, Lys 57 and Arg 58. The appropriate amino-acid substitutions were made using the program QUANTA (version 3.0, Polygen Corp). The geometry



of the calcium-binding loop (residues 28–34) was adjusted to more closely resemble that found in the structure of the bovine pancreatic enzyme. This hybrid model was then refined using the molecular dynamics program AMBER^{24,25} (version 3.0A), while holding all alanine, cysteine and glycine residues fixed. Starting with this model as the search structure, solutions for the rotation and translation functions were obtained using the program XPLO^{26,27} (version 2.1). The structure was refined alternately with XPLO^{26,27} and with the program PROLSQ²⁸. The quality of the model was analysed and improved at intervals during the refinement using the molecular graphics program FRODO²⁹. The present crystallographic R factor is 0.178 for all data between 6.0 and 2.2 Å, with tight restraints on the molecular geometry. There were 77 water molecules included in the last least-squares refinement. On average, bond lengths, interbond angle distances and planarities deviate from ideal by 0.017, 0.029 and 0.034 Å, respectively.

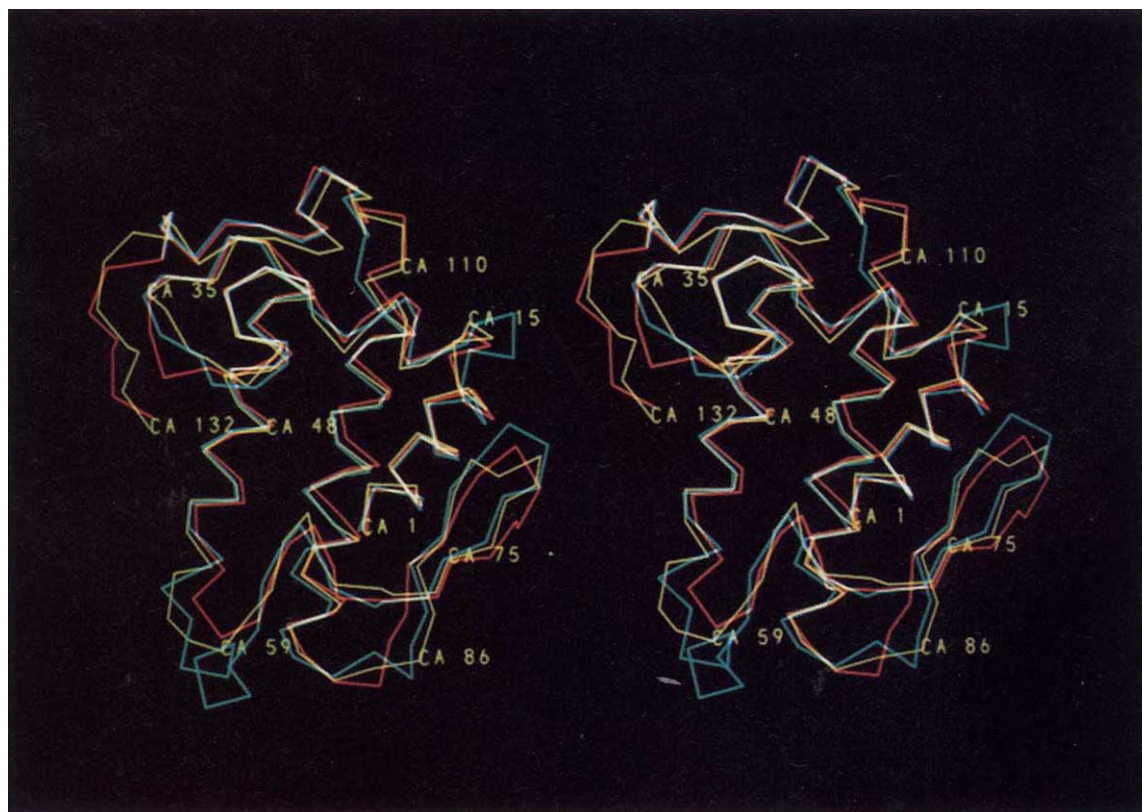
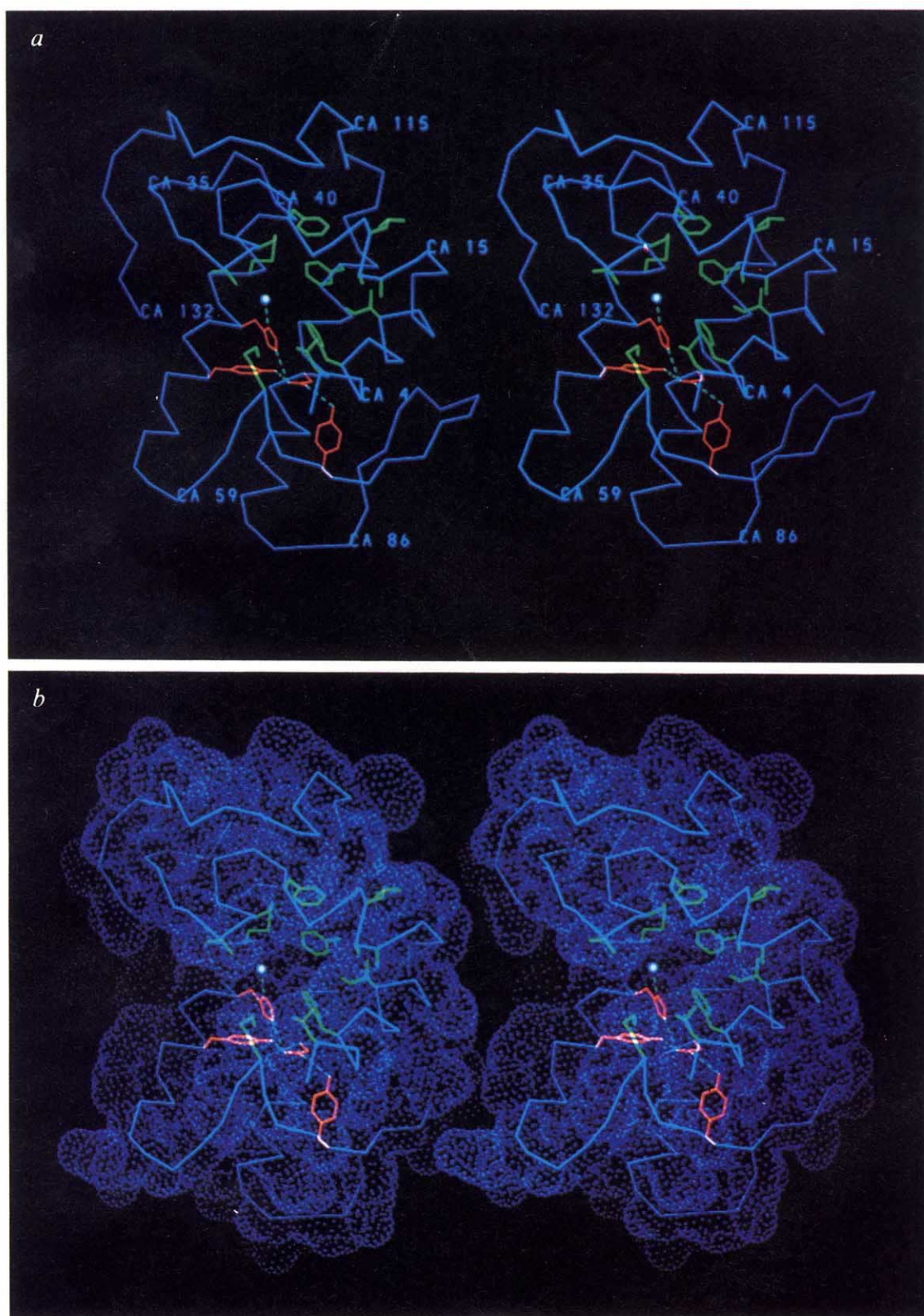


FIG. 2 Stereoview of the α -carbon backbone of human s-PLA₂ (yellow) superimposed on the α -carbon backbones of a group I, bovine (blue), and a group II, *C. atrox* (red), PLA₂.

FIG. 3 Stereoview of the functionally important residues in human s-PLA₂. *a*, An α -carbon tracing of the enzyme is shown in blue. The conserved catalytic residues, His 48 and Asp 99, and the invariant residues, Tyr 52 and Tyr 73, are shown in red; they are connected through a network of hydrogen bonds to the active site water molecule (in blue). Shown in green are the side chains of the residues forming the hydrophobic channel (Leu 2, Phe 5, Ile 9, Ala 18, Leu 19, Cys 29, Val 31, Cys 45, Lys 69 and Phe 106) and part of the interfacial binding surface (Ala 17 and Phe 24). *b*, The α -carbon tracing with side chains as above, also showing the van der Waals surface in blue. The section shown is from the base of the pocket up to the putative interfacial surface. Note that the hydrophobic residues line a deep pocket in the centre of the molecule that ends near the active site water molecule Wat 218.



the five equatorial ligands could be contributed by the two carboxylate oxygens of Asp 49, the main-chain carbonyl of Gly 32, and a water molecule, Wat 242, which is slightly displaced toward the main-chain carbonyl of His 28. By an easily accommodated rotation of Gly 30 in the calcium-binding loop, its main-chain carbonyl could satisfy the fifth position. The His 28 carbonyl is in one axial position, and only the last axial ligand is completely missing. A comparison of the human s-PLA₂

structure with that of the bovine pancreatic enzyme, which does contain calcium, shows a close similarity in their calcium-binding regions (Fig. 2). Although calcium is required for both substrate binding and catalysis, its absence in our crystal produces only small structural differences. The structural details surrounding the expected calcium-binding site of human s-PLA₂ are therefore likely to be relevant to substrate binding and catalysis.

The details of the PLA₂ catalytic mechanism have been revealed by structural studies of complexes of the enzyme with phosphonate transition-state analogues^{3,18,19}. It is instructive to examine details of our structure in light of these structures and their implied mechanisms. The functionally important and invariant residues, His 48, Tyr 52, Tyr 73 and Asp 99 from human s-PLA₂, are virtually superimposable on those of both bovine pancreatic and *C. atrox* PLA₂. Hence the hydrogen-bonding network that connects these residues remains intact (Fig. 3a). One functionally very important water molecule, Wat 218, is hydrogen-bonded to Nδ1 of His 48 (Figs 1 and 3a), as in all high resolution PLA₂ structures that lack inhibitor. This water molecule has been proposed to be the attacking nucleophile^{3,4} and is thought to become part of the tetrahedral intermediate formed during catalysis. A second water molecule, Wat 242, which we have identified as one of the apparent equatorial Ca²⁺ ligands, lies about 2.8 Å from Wat 218, toward the vacant Ca²⁺ site in our structure. According to the proposed mechanism, this water molecule is replaced by the oxyanion of the tetrahedral intermediate; the oxyanion then becomes a ligand of the calcium ion which the latter helps stabilize. Thus, Wat 218, and probably Wat 242, are the counterparts of two water molecules seen in the *Naja naja atra* structure (see Fig. 4 of ref. 3) and are important reference points for an understanding of the expected binding of substrate and formation of the tetrahedral intermediate in human s-PLA₂.

A hydrophobic channel has been clearly identified in the structures of PLA₂ complexes with inhibitors^{18–20}. A corresponding region is found in human s-PLA₂ (Fig. 3) and appears as a deep pocket which is lined, as in other PLA₂s, with the invariant hydrophobic residues Phe 106 at the base of the pocket and Leu 2, Phe 5 and Ile 9 on the N-terminal helix. Residues which line the pocket, but which are infrequently found at these positions in other PLA₂s, include Ala 17, Phe 24 and Val 31; they are distributed on the opposite side of the pocket. Additional residues which line the hydrophobic pocket in our enzyme include Ala 18, Leu 19, Cys 29 and Cys 45, the latter two forming a disulphide bond. Lys 69 is positioned on one side of the pocket, where it could serve its proposed dual function of moveable hydrophobic 'flap' and substrate ligand¹⁹. The hydrophobic channel is narrower in the human enzyme than in bovine pancreatic or *C. atrox* PLA₂. This narrowing is caused by displacement of residues 20–24 by about 0.5 Å into the channel and by the presence of His 6, a residue which has been observed at this position only in the human enzyme²¹. The latter residue intrudes more deeply into the channel than the more flexible and often charged residues, such as lysine, which are usually found at this position. This suggests that compounds with smaller cross-sections may be better inhibitors.

Human s-PLA₂ is a very positively charged molecule with 23 arginines and lysines but only eight glutamic and aspartic acid residues, with two of the latter, Asp 49 and Asp 99, required for function. Several of the basic residues encircle the hydrophobic channel and probably account for this enzyme's striking preference for negatively charged substrates, such as phospholipids found in *Escherichia coli* membranes. Two positively charged residues, Arg 7 and Lys 10, are particularly close to the channel and lie exposed along the N-terminal helix. One negatively charged residue, Glu 16, is also near the hydrophobic channel, whereas the remainder are on the opposite surface of the enzyme, away from the channel. It is not clear why there is such an overall excess of positively charged residues. As the distribution of such charges is fairly uniform they may enhance nonspecific association with negatively charged bacterial membranes which may in turn aid binding and catalysis.

Binding of phosphonate and amide phospholipid analogues to other PLA₂s produces no appreciable structural changes^{3,20} in the active-site region. The X-ray structure of the uncomplexed human s-PLA₂ presented here is probably the best model currently available for structure-based drug design. □

Received 4 April; accepted 17 May 1991.

- Davidson, F. F. & Dennis, E. A. *J. molec. Evol.* **31**, 228–238 (1990).
- Waite, M. in *Phospholipases* (ed. Hanahan, D. J.) 155–241 (Plenum, New York, 1987).
- Scott, D. L. *et al. Science* **250**, 1541–1546 (1990).
- Verheij, H. M. *et al. Biochemistry* **19**, 743–750 (1980).
- Vadas, P., Pruzanski, W. & Stefanski, E. *Agents & Actions* **24**, 320–325 (1988).
- Vadas, P. & Pruzanski, W. *Lab Invest.* **4**, 391–404 (1986).
- Chang, J., Musser, J. H. & McGregor, H. *Biochem. Pharmacol.* **36**, 2429–2436 (1987).
- Forst, S. *et al. Biochemistry* **25**, 8381–8385 (1986).
- Chang, H. W., Kudo, I., Tomita, M. & Inoue, K. *J. Biochem., Tokyo* **102**, 147–154 (1987).
- Stefanski, E., Pruzanski, W., Sternby, B. & Vadas, P. *J. Biochem., Tokyo* **100**, 1297–1303 (1986).
- Kramer, R. M. *et al. J. biol. Chem.* **264**, 5768–5775 (1989).
- Seilhamer, J. J. *et al. J. biol. Chem.* **264**, 5335–5338 (1989).
- Fleer, E. A. M., Verheij, H. M. & deHaas, G. H. *Eur. J. Biochem.* **82**, 261–269 (1978).
- Randolph, A. & Heinrikson, R. L. *J. biol. Chem.* **257**, 2155–2161 (1982).
- Kramer, R. M., Johansen, B., Hession, C. & Pepinsky, R. B. *Adv. exp. med. Biol.* **275**, 35–54 (1990).
- Dijkstra, B. W., Kalk, K. H., Hol, W. G. J. & Drenth, J. *J. molec. Biol.* **147**, 97–123 (1981).
- Brunie, S., Bolin, J., Gewirth, D. & Sigler, P. B. *J. biol. Chem.* **260**, 9742–9749 (1985).
- White, S. P., Scott, D. L., Otwinowski, Z., Gelb, M. H. & Sigler, P. B. *Science* **250**, 1560–1563 (1990).
- Scott, D. L., Otwinowski, Z., Gelb, M. H. & Sigler, P. B. *Science* **250**, 1563–1566 (1990).
- Thunnissen, M. M. G. M. *et al. Nature* **347**, 689–691 (1990).
- Heinrikson, R. L. *Meth. Enzym.* **197**, 210–214 (1991).
- Grinnell, B. W., Berg, D. T., Walls, J. & Yan, S. B. *Biotechnology* **5**, 1189–1192 (1987).
- Howard, A. J. *et al. J. appl. Crystallogr.* **20**, 383–387 (1987).
- Weiner, S. J. *et al. J. Am. chem. Soc.* **106**, 765–784 (1984).
- Weiner, S. J., Kollman, P. A., Nguyen, D. T. & Case, D. A. *J. comput. Chem.* **7**, 230–252 (1986).
- Brünger, A. T., Kuriyan, J. & Karplus, M. *Science* **235**, 458 (1987).
- Brünger, A. T. *Acta crystallogr.* **A46**, 46–57 (1990).
- Hendrickson, W. A. & Konner, J. H. in *Biomolecular Structure, Function, Conformation and Evolution* (ed. Srinivasan, R.) **1**, 43–57 (Pergamon, Oxford, 1981).
- Jones, T. A. *J. appl. Crystallogr.* **11**, 268–272 (1978).

Insect paralysis by baculovirus-mediated expression of a mite neurotoxin gene

Michael D. Tomalski & Lois K. Miller

Departments of Entomology and Genetics, University of Georgia, Athens, Georgia 30602, USA

FEMALE mites of the species *Pyemotes tritici* inject an extremely potent venom into their insect prey that causes muscle-contraction and paralysis^{1–4}. These mites are able to paralyze insects 150,000 times their size⁵ and their venom is effective in a broad range of insect species⁴. A toxin (TxP-I) associated with the mite venom apparatus causes immediate muscle-contraction paralysis when injected into insects but not mice⁶. In this report, we describe the cloning, sequencing and expression of a complementary DNA (Tox-34) encoding TxP-I. Insect cells infected with a recombinant baculovirus (vEV-Tox34) expressing Tox-34 secrete three polypeptides related to TxP-I which cause paralysis on injection. Larvae infected with vEV-Tox34 become paralyzed during infection, thus reflecting the potential application of this toxin gene in insect biocontrol methods. The toxin gene expression system will also allow further exploration of the neurophysiological basis of its insect-specific effects.

A toxin-encoding DNA was isolated from a cDNA expression library prepared from 20 µg polyadenylated messenger RNA⁷ isolated from a mixture of gravid and host-seeking female mites. The λZAPII (Stratagene) library was screened with a polyclonal anti-TxP-I antiserum raised against homogeneous TxP-I purified from mites⁶. The cDNA inserts of selected immunoreactive phage plaques were subcloned into plasmids and analysed by hybridization⁸ to degenerate oligodeoxynucleotide probes deduced from the N-terminal TxP-I sequence⁶ and by nucleotide sequencing⁹. A degenerate 62-residue-long probe hybridized to one immunologically positive clone at a site that includes the N terminus of TxP-I.

The nucleotide sequence of the *EcoRI* fragment containing this cDNA, known as Tox34, contains an open reading frame which could encode a polypeptide of 291 amino acids of relative molecular mass 33,000 (*M_r* 33K) (Fig. 1). The predicted polypeptide includes a 21-amino-acid sequence (residues 40–60) which matches the empirically determined N-terminal sequence⁶

raised to purified TxP-I or with preimmune serum (Fig. 2b). Three polypeptides in vEV-Tox34 cell lysates, similar in size to TxP-I and TxP-II, reacted specifically with the antibody, indicating that these three polypeptides are indeed related to TxP-I and TxP-II. (Immunoreactive bands below the TxP-I band in cell lysates are probably degradation products.) A single cross-reactive protein, similar in size to TxP-I, was found in the extracellular media (Fig. 2b, lane S) indicating that this product is secreted, stable TxP-I. But only a single TxP-I-related product was observed in the western blot of the extracellular fraction, even though the two larger polypeptides were detected by pulse-labelling in this fraction (Fig. 2a). This suggests that the TxP-II products are unstable in the extracellular fluid and are perhaps precursors of TxP-I.

Although TxP-I seems to be the mature form of the toxin, the precise nature of the two TxP-II precursor forms is not known. TxP-I could be synthesised as a 291- or 272-amino-acid polypeptide precursor, depending on whether translation initiates at Met 1 or 13, respectively (see Fig. 1). Post-translational processing between Ile 39 and Asp 40 would then generate the

single, mature TxP-I. Alternatively, two cleavages of a single preprotoxin that contains signal and protoxin segments would also result in TxP-I.

To determine if the secreted Tox-34 products were toxic, 5 μ l of extracellular fluid from wild-type, vEV-Tox34, or mock-infected insect cells were injected into larvae of the wax moth *Galleria mellonella*. Injection of the extracellular fluid from vEV-Tox34-infected cells produced a fast, muscle-contractive paralysis within 2–5 min. The symptomatology of paralysis was identical to that observed in larvae injected with TxP-I purified from mites⁶. With lower doses of toxin, insects that do not exhibit immediate muscle-contractive paralysis show no signs of paralysis later. Larvae injected with fluid from either mock-infected or wild-type-infected cells showed no paralytic symptoms. Thus, Tox-34 products were biologically active and produced paralytic symptoms identical to those of purified mite toxins. The toxin, however, elicited no toxic symptoms when fed to neonates or to late instar larvae.

Insects infected with recombinant viruses expressing Tox-34 responded more rapidly to infection than insects infected with wild-type virus. Fifth instar *Trichoplusia ni* larvae injected with a lethal dose of vEV-Tox34 in the budded form were paralysed within two days of infection, whereas larvae injected with the same dose of wild-type (wt) virus never showed symptoms (Table 1). Eventually, all virus-injected larvae died of a typical viral infection, showing that paralysis does not prevent the replication of vEV-Tox34 in larvae. We also constructed an occlusion-positive recombinant virus, vSp-Tox34 (M.D.T. and L.K.M., manuscript in preparation), expressing Tox-34 under the control of a strong very late promoter, P_{SynXIV} (ref. 13). On oral infection with a low dose of occluded virus, neonate larvae infected with vSp-Tox34 succumbed to paralysis and/or death at a faster rate than those infected with wild-type virus (Fig. 3). Thus, recombinant viruses expressing Tox-34 produce biologically active toxin and induce an impressive paralytic response during infection.

The successful cloning and expression of Tox-34 constitutes a major step in characterizing and exploiting this new class of

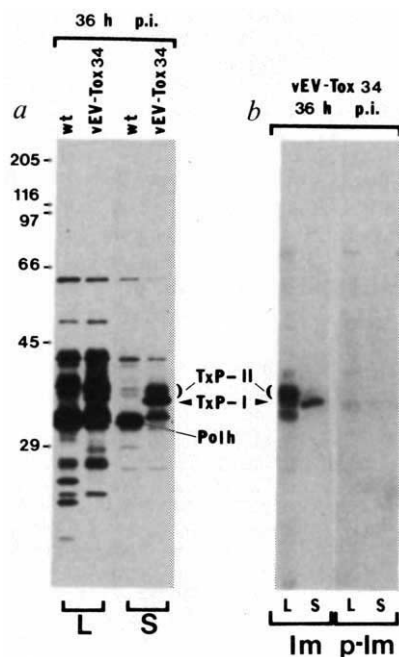


FIG. 2 Expression of Tox-34 in baculovirus-infected insect cells. *a*, An autoradiograph of an SDS-polyacrylamide gel containing proteins from lysates (L) or extracellular fluid (S) of cells infected with wild-type (wt)-AcMNPV or vEV-Tox34 and radiolabelled at 36 h post-infection (p.i.). Positions of M_r markers ($M_r \times 10^{-3}$) are shown on the left. The gel positions of authentic TxP-I and TxP-II are shown to the right of the panel. The position of the polyhedrin protein (polh) is also indicated; owing to the fragility of cells at 36 h p.i., low levels of this abundant protein are found in the supernatants. *b*, Western blots of proteins from cell lysates (L) or extracellular fluid of vEV-Tox34-infected cells were probed with anti-TxP-I antibody (Im) or preimmune serum (p-Im).

METHODS. *S. frugiperda* cells were infected with wt-AcMNPV or vEV-Tox34 as described¹². Thirty-six hours p.i., the cells were pulse-labelled (1 h) with [³⁵S]cysteine following a 1 h incubation in cysteine-free media. The cells were washed three times with phosphate-buffered saline to remove unincorporated radioisotope and overlaid with unsupplemented TC-100 medium. After a 2-h incubation, the extracellular fluid and the cells were collected separately. Proteins from cell lysates and extracellular fluid were suspended in a buffer containing SDS and DTT, and analysed by SDS-PAGE¹⁷. For western blotting¹⁸, proteins were electrophoretically transferred to Immobilon membranes (Millipore) and incubated with either preimmune or immune serum raised against purified TxP-I (ref. 6). Immunopositive proteins were visualized with alkaline phosphatase-conjugated goat antirabbit IgG and nitroblue tetrazolium (Sigma)¹⁹. Proteins from mock-infected cells did not exhibit any detectable reaction with the antibody and were not included in the figure.

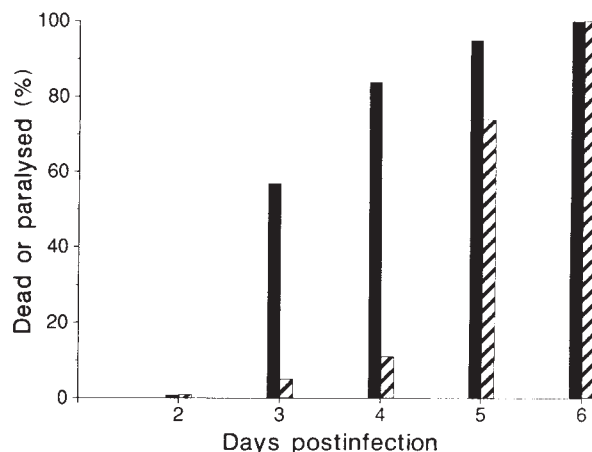


FIG. 3 Effects of Tox-34 expression after oral infection of larvae with occluded viruses. *T. ni* neonate larvae (<12 h old) were infected with occluded virus of wt-AcMNPV (slashed bars) or vSp-Tox34 (solid bars), a recombinant virus expressing both Tox-34 and polyhedrin genes. The neonates were placed on diet containing 1.6×10^4 polyhedra ml^{-1} of diet, a concentration approximating an LC_{50} dose. After 24 h (day 1), the larvae were transferred individually to fresh diet and scored daily for paralysis or death. Ninety larvae were used for each virus and for an uninfected larval control. The mortality was 4% in the uninfected control larvae. Although the experiment was continued for 10 days, the number of insects that had died remained constant after 6 days; surviving insects were not infected using this low virus dose. The percentage dead or paralysed larvae presented in the graph refer only to those insects that eventually died from virus infection; of the 90 larvae exposed to each virus, 34 larvae were infected with vSp-Tox34 and 50 larvae were infected with wild-type virus.

insect toxins. Not only does this work provide a convenient source of toxin for future biochemical and neurophysiological studies, it also opens many new avenues for the development of biological control agents. Tox-34 is the first toxin gene found to increase greatly the rate at which a baculovirus controls its host. These viruses are used as highly specific pest control agents but improvements in the rate at which the viruses control host-feeding damage are needed for their more generalized use in agriculture¹⁴⁻¹⁵. Tox-34 will also be of much interest in the development of other biorational and biological strategies of insect pest control. □

Received 18 February; accepted 13 May 1991.

- Krczal, H. *Acarina* **1**, 385-625 (1959).
- Moser, J. C. *Trans. R. Ent. Soc. London* **127**, 185-191 (1975).
- Vaivaniukul, P. & Haramoto, F.H. *Proc. Hawaii Ent. Soc.* **20**, 443-445 (1968).
- Weiser, J. & Slama, K. *Ann. Ent. Soc. Am.* **57**, 479-482 (1964).
- Tomalski, M. D. *et al. Toxicon* **26**, 127-132 (1988).
- Tomalski, M. D. *et al. Toxicon* **27**, 1151-1167 (1989).
- Davis, L. G., Dübner, M. D. & Battey, J. F. *Basic Methods in Molecular Biology* (Elsevier, New York, 1986).
- Tomalski, M. D., Wu, J. & Miller, L. K. *Virology* **167**, 591-600 (1988).
- Sanger, F., Nicklen, S. & Coulson, A. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463-5467 (1977).
- Miller, L. K. *A. Rev. Microbiol.* **42**, 177-179 (1988).
- Ooi, B. G., Rankin, C. & Miller, L. K. *J. molec. Biol.* **210**, 721-736 (1989).
- Miller, D. W., Safer, P. & Miller, L. K. in *Genetic Engineering: Principles and Methods* Vol. 8 (eds Setlow, J. & Hollaender, A.) 277-298 (Plenum, New York, 1986).
- Wang, X., Ooi, B. G. & Miller, L. K. *Gene* **100**, 131-137 (1991).
- Miller, L. K., Lingg, A. J. & Bulla, L. A. Jr. *Science* **219**, 715-721 (1983).
- Carbonell, L. F. *et al. Gene* **73**, 409-418 (1988).
- Bishop, D. H. L. *et al. in The Release of Genetically Engineered Micro-organisms* (eds Sussman, M., Collins, C. H., Skinner, F. A. & Skinner, D. E.) 143-179 (Academic, New York, 1988).
- Laemmli, U. K. *Nature* **227**, 680-685 (1970).
- Towbin, H., Staehelin, R. & Gordon, J. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4350-4354 (1979).
- Blake, M. S. *et al. Anal. Biochem.* **136**, 175-179 (1984).

ACKNOWLEDGEMENTS. We thank W. A. Bruce (USDA-ARS Bee Research Lab, Beltsville, Maryland) for providing *P. tritici*, S. Hancher (Abbott Labs, Chicago, Illinois) for *T. ni* eggs, J. Todd and S. Hilliard for expert technical assistance, and D. O'Reilly, N. Giles, S. Kushner, J. Willis, M. Blum and A. Lea for reading the manuscript. This work was supported in part by the National Institute of Neurological and Communicative Disorders and Stroke.

Construction of an improved baculovirus insecticide containing an insect-specific toxin gene

Lorna M. D. Stewart, Mark Hirst, Miguel López Ferber, Alison T. Merryweather, P. Jane Cayley* & Robert D. Possee

NERC Institute of Virology and Environmental Microbiology, Mansfield Road, Oxford OX1 3SR, UK

* Wellcome Environmental Health, McIntyre House, High Street, Berkhamsted, Hertfordshire HP4 2DY, UK

BACULOVIRUSES provide alternatives to chemicals for controlling insect pests¹⁻⁴ and can be applied by spraying. Baculoviruses have a limited host range, but work relatively slowly. They are dissolved in the midgut of insect larvae to release infectious virions which enter gut epithelial cells and begin to replicate. Replication in other organs causes extensive tissue damage and eventually death. This process can take 4-5 days, but in the field may last for more than a week, allowing the larvae to feed for longer and thereby damaging the host plant. Baculovirus expression vectors expressing foreign genes^{5,6}, such as those for insect-specific toxins, hormones or enzymes, might alleviate this problem⁷⁻¹¹. We have now constructed a recombinant baculovirus derived from *Autographa californica* nuclear polyhedrosis virus containing an insect-specific neurotoxin from the venom of the North African (Algerian) scorpion, *Androctonus australis* Hector¹². The neurotoxin acts by causing specific modifications to the Na⁺ conductance of neurons, producing a presynaptic excitatory effect leading to paralysis and death^{15,16}; it has no effect in mice^{17,18}.

Expression of the neurotoxin by the virus causes a reduction in the time required to kill the host insect.

In most baculovirus expression vectors the polyhedrin coding region is replaced by the foreign DNA producing unstable, non-occluded virions. We preserved the occlusion of the virus by employing a transfer vector pAcUW2B (ref. 19) which expresses foreign genes using the p10 promoter inserted upstream of the intact polyhedrin gene. Two polyhedrin-positive viruses were constructed (Fig. 1a), with the *A. australis* Hector insect toxin (AaHIT) coding region, made from synthetic oligonucleotides (Fig. 1b), under the control of the *A. californica* nuclear polyhedrosis virus (AcNPV) p10 promoter. In one version, AcST-1, this sequence began with a simple translation initiation codon. In the other, AcST-3, a copy of the AcNPV gp67 signal peptide sequence²⁰ was incorporated at the 5' end of the coding region, to facilitate protein secretion. The first virus was prepared as a control to ensure that the predicted effect of the toxin was due to secretion of the toxin rather than simple leakage from the cells.

The virus-directed intracellular expression of the toxin gene was demonstrated by examination of proteins from *Spodoptera frugiperda* cells infected with AcST-3 after pulse labelling with [³⁵S]cysteine (Fig. 2a). Synthesis of the putative AaHIT was first detected 18 h after infection, maximal expression occurred at about 36 h, and was followed by a decrease at 48 h. Other proteins are also produced in lesser quantities at this time. Similar results were obtained with AcST-1 (data not shown). Consistent detection of the secreted form of the toxin in medium from AcST-3-infected cells was problematic because of the very

TABLE 1 LD₅₀ (polyhedra per larvae) and ST₅₀ (h) values for AcST-3 and AcNPV in *T. ni* larvae

Virus	LD ₅₀	ST ₅₀
AcNPV	44 (35-55)*	113.1 (112-115)
AcST-3	31 (27-37)†	85.8 (85-87)‡

Lethal dose (LD₅₀) measurements: Fifty 2nd-instar larvae (about 0.6-0.7 mg) were individually fed five serial dilutions of each virus stock on a small plug of artificial diet in a microtitre plate. The maximum dose of the virus should ideally give over 90% mortality, and the minimum dose should give 10% mortality. Nonengineered strains of AcNPV fulfilled these conditions with 120, 60, 30, 15 and 7.5 polyhedra per larva. Larvae that consumed the dose in 24 h were transferred to individual containers of artificial diet and examined daily. Cadavers were removed, smeared on a slide and the cause of death confirmed. The data were analysed using probit analysis to determine the LD₅₀ values²⁷. Survival time (ST₅₀) measurements: To produce the neonate larvae used in these studies, adults were reared in cages containing filter papers for oviposition. The filters carrying the eggs were surface sterilized and retained in plastic containers. After hatching, the neonates were starved for 3-6 h before droplet feeding with each virus suspension (2 × 10⁶ polyhedra ml⁻¹). The suspensions were coloured with 5% blue food dye to allow visualization of feeding. Small droplets of virus were placed on a Petri dish in concentric rings. The larvae were put in the centre of the rings, after which they moved through the droplets, taking in a small volume of liquid before crawling on to the lid of the dish. Larvae which were damaged in handling remained in the centre, allowing synchronous feeding and the selection of healthy individuals. Previous work showed that in these type of experiments, larvae of *T. ni* consume 0.0087 (±0.0023) μl^{28,29}. After feeding, the larvae were maintained in individual containers with artificial diet at 23 °C. After 24 h, the larvae killed by handling were removed. Thereafter the larvae were checked at frequent intervals. Dead larvae were removed and the cause of death diagnosed by appearance and microscopic examination. ST₅₀ calculations were made with the Vistat program (Boyce Thompson Institute, Ithaca, New York, 1990). Polyhedra used in the LD₅₀ and ST₅₀ tests had been purified from virus-infected *T. ni* larvae in the presence of 0.1% SDS using differential centrifugation and sucrose gradients³⁰.

* 95% confidence limits are indicated within brackets after each value.

† A statistically significant reduction in LD₅₀ relative to the wild-type virus ($\alpha=0.05$).

‡ A statistically significant reduction in ST₅₀ relative to the wild-type virus ($\alpha=0.0001$).

low levels that were apparently released, and spontaneous lysis of the virus-infected cells late in infection. Therefore, secretion was assessed by HPLC purification of the toxin from the media of AcST-3-infected *S. frugiperda* cells. This revealed a peak in the optical density readings at 280 nm around fractions 26–28, which was also present in the total venom but was absent in cells infected with wild-type virus or AcST-1 (data not shown). The concentration of AaHIT in AcST-3-infected cell culture medium was estimated to be 400 ng ml⁻¹ from the HPLC trace. Analysis of the fractions by immunoblot (with guinea-pig serum raised against a synthetic AaHIT peptide) confirmed the presence of the toxin (data not shown).

To determine whether the secreted protein was biologically active, the appropriate HPLC fractions were injected into *Musca domestica* adults, which were chosen because previous studies had shown that they were more sensitive to the toxin than Lepidopteran species¹⁸. Injection of 300 ng of the purified AaHIT into fourth instar *Trichoplusia ni* larvae (18 ± 3 mg) pro-

duced no response 1 h after injection when compared with controls injected with PBS medium. By contrast, 5 ng toxin could paralyse or kill 50% of a group of *M. domestica* adults (data not shown). The HPLC-purified AaHIT from AcST-3-infected cell culture medium demonstrated a biological activity which was similar to that of the natural toxin. Injection of the toxin-containing fractions resulted in paralysis of insets in 3 h followed by death in 24 h (Fig. 2b). For fraction 26 (AcST-3) in Fig. 2b, this represented a toxin dose of about 64 ng. The apparent size of the virus-encoded toxin in polyacrylamide gels was similar to the AaHIT component of natural venom, therefore the shift in fraction activity may be accounted for by sample collection variability from the reverse-phase column.

The biological activity of the recombinant viruses was first assessed by feeding purified polyhedra to third instar *T. ni* larvae. The pathology associated with AcST-3 infection differed from that with AcST-1 or AcNPV. Infection with AcST-1 or AcNPV resulted in massive viral-induced tissue damage leading to

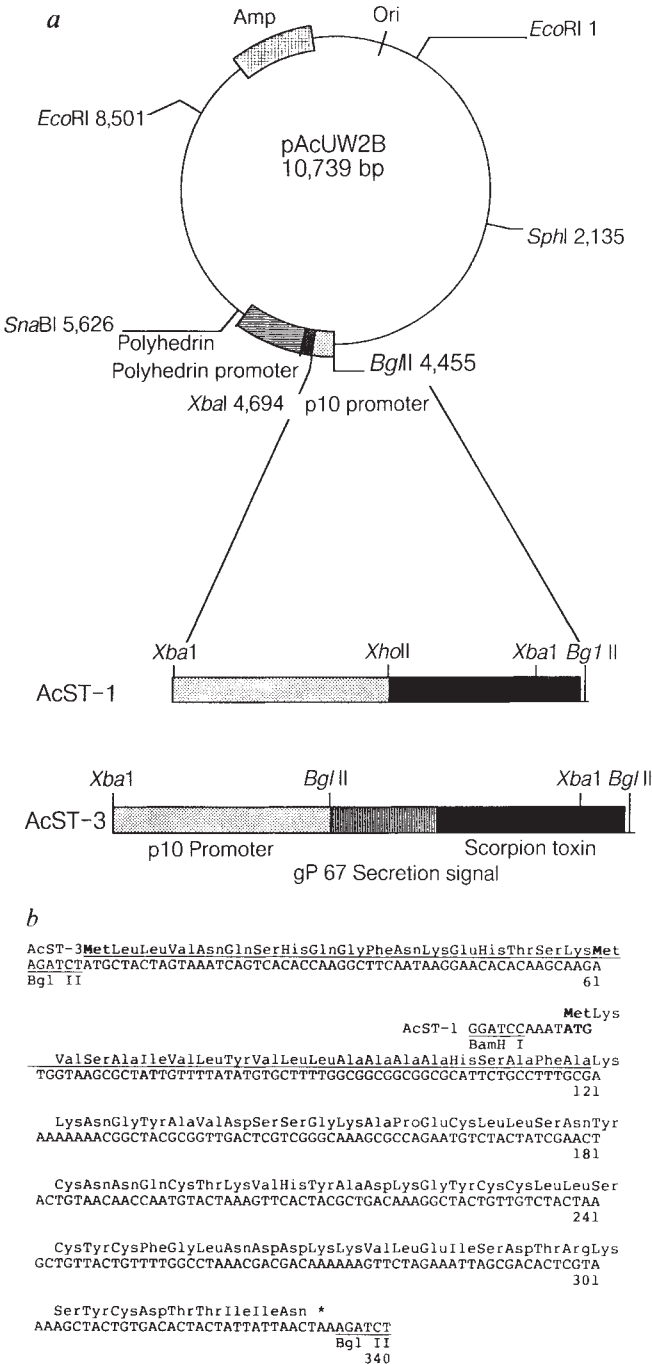


FIG. 1 Construction of recombinant viruses. *a*, Construction of baculovirus transfer vectors. The AaHIT coding regions for AcST-1 (no signal peptide) and AcST-3 (with AcNPV gp67 signal peptide) were inserted into the transfer vector pAcUW2B¹⁹ under the control of the p10 promoter. This vector preserves the complete polyhedrin gene. *b*, Sequence of the synthetic AaHIT coding region. The sequence from 121–331 represents the coding region of the AaHIT. In AcST-1 this is preceded by a translation initiation codon and the four nucleotides normally located before the AcNPV polyhedrin ATG (ref. 21). In AcST-3 the translation initiation codon was replaced by a copy of the putative signal peptide region (underlined) from the AcNPV gp67 gene²⁰. The two potential ATG codons in this sequence are highlighted in bold text. Flanking restriction enzyme sites *Bam*HI and *Bgl*II are indicated. METHODS. The synthetic DNA encoding the AaHIT was constructed by the insertion of four complementary pairs of oligonucleotides into the polycloning site of a modified pUC18 vector containing a *Bgl*II site. For AcST-3 the synthetic gp67 signal peptide coding region²⁰ was inserted at the *Bam*HI site upstream of the ATG codon in AcST-1 and a favourable protease recognition site constructed by using site-directed mutagenesis²² after subcloning in pAcCL29²³ to provide a source of single-stranded DNA. The transfer vectors pAcST-1 and pAcST-3 were constructed by subcloning the appropriate coding regions into pAcUW2B¹⁹. Recombinant viruses were produced by mixing pAcST-1 or pAcST-3 with linearized, polyhedrin-negative AcRP6-SC DNA²⁴ and cotransfecting *S. frugiperda* cells in the presence of lipofectin (GIBCO-BRL). Subsequent plaque purification of the progeny viruses identified polyhedrin-positive recombinants that were repurified to homogeneity. The genomic structures of the recombinant viruses were confirmed by Southern hybridization of DNA purified from virus-infected cells²⁵.

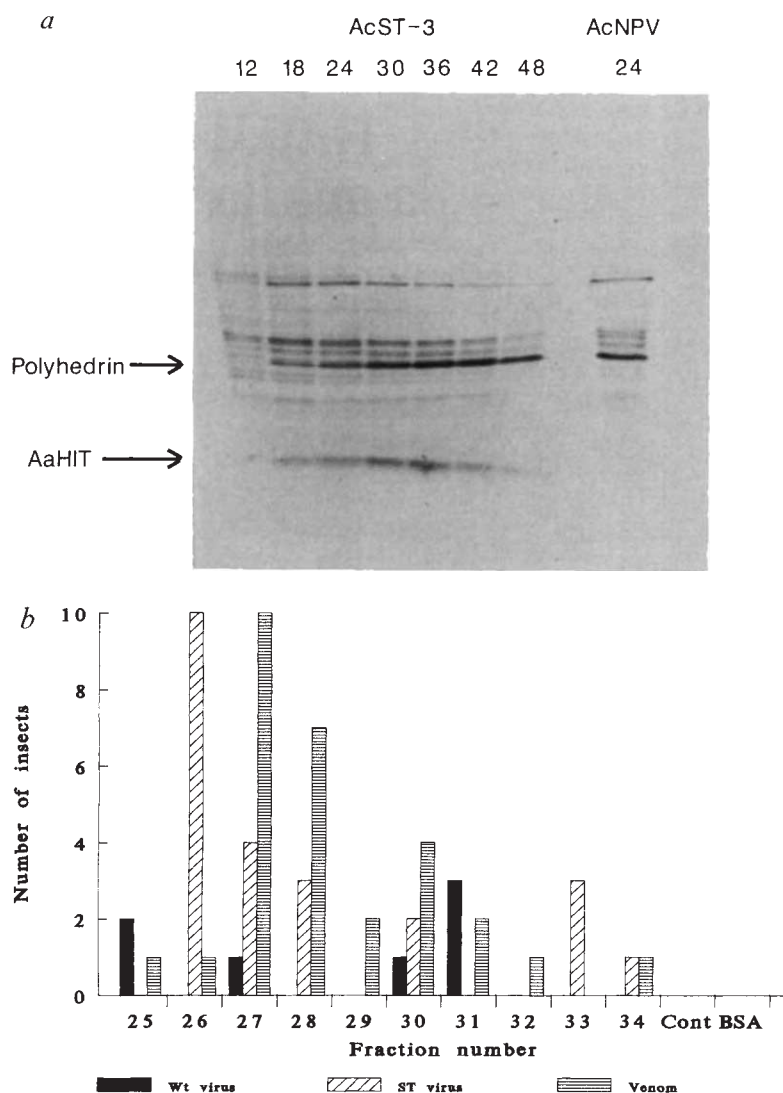


FIG. 2 Synthesis of recombinant AaHIT and assay of its biological activity. **a**, Intracellular, radiolabelled proteins in *S. frugiperda* cells infected with AcST-3 or AcNPV. Infected cells were pulse-labelled with [35 S]cysteine at the times after infection (h) indicated above each track. Polyhedrin and AaHIT are indicated. Similar results were obtained for AcST-1. **b**, Biological assay of secreted AaHIT. Total crude AaH venom (Venom), the medium from *S. frugiperda* cultures infected with AcNPV (Wt virus) or the medium from cultures infected with AcST-3 (ST virus) were purified using HPLC and each fraction (x axis) assayed by injecting 10 *M. domestica* adults and recording the number of insects paralysed or dead after 24 h (y axis). To kill or paralyse every insect required 64 ng AaHIT (calculated from previous calibration experiments).

METHODS. The *S. frugiperda* cells infected with virus (multiplicity of infection of 10 PFU per cell) were radiolabelled with [35 S]cysteine for 1 h as previously described⁹. Protein extracts were fractionated in a denaturing reducing 10–30% polyacrylamide gel²⁶ and detected using autoradiography. For HPLC analysis the medium from AcST-3-infected cells (2×10^6) was concentrated fourfold using the Centricon filtration system (3K) by centrifugation at 5,000g for 3 h on a Sorval RC5 centrifuge. Purification was carried out by loading 100 μ l concentrated material onto the HPLC column using an LCD system with a Spectroflow 738 detector set at 280 nm. A Spherisorb, C8 25 cm column was used for separation. Solvent A was 0.15 M ammonium formate, pH 2.7; solvent B was acetonitrile. A linear gradient was run from 15% to 40% B in A over 40 min at a flow rate of 1 ml min⁻¹ and 40 $\times$ 1 ml fractions were collected. The fractions were freeze-dried to remove acetonitrile and ammonium formate, then were resuspended in 100 μ l of 0.1 mg ml⁻¹ BSA in water. Total crude scorpion venom was obtained from Sigma as a freeze-dried sample (1 mg), resuspended in 20 μ l of dimethylsulphoxide, then diluted to 100 μ l in water. Precipitated material was removed by centrifugation at 15,000g for 5 min. Purification of 20 μ l was carried out by HPLC as before, but the freeze-dried sample was resuspended in 200 μ l 0.1 mg ml⁻¹ BSA in water. Samples (4 μ l) of each fraction (natural AaHIT- or AcST-3-derived material) were injected into the dorso lateral area of the abdomens of 10 adult *M. domestica* (male and female) using a 30-gauge syringe.

liquifaction of the larvae producing a creamy coloration. By contrast, infection of larvae with AcST-3 resulted in larvae which maintained their structural integrity and remained green.

The improved insecticidal activity of the recombinant AcST-3 was demonstrated by reductions in both median survival time (ST₅₀) and median lethal dose (LD₅₀) values in *T. ni* larvae (Table 1). The ST₅₀ value for AcST-3 was reduced to 86 h compared with 113 h for AcNPV, a decrease of 25%. The virus dose used was about 17 polyhedra per larva. To investigate

whether infection with AcST-3 could reduce crop damage by the larvae, feeding trials on cabbage leaves (*Brassica* sp.) were performed. Third instar *T. ni* larvae were fed a small plug of semisynthetic diet contaminated with polyhedra and then placed onto cabbage leaves for continued feeding. Insects given the AcST-3 recombinant virus consistently caused less damage to the leaves than individuals infected with AcNPV (Fig. 3). This difference was calculated to be 50% in later experiments where leaf area was measured before and after feeding to virus-infected insects.

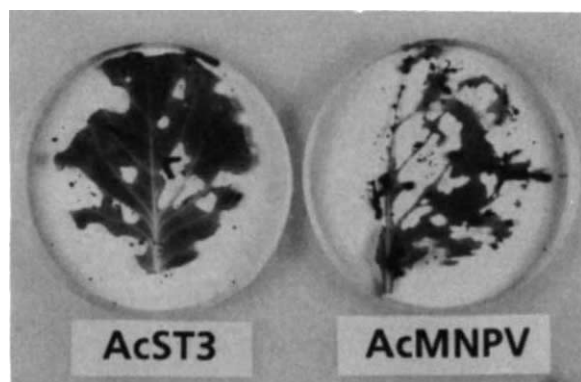


FIG. 3 Reduction in feeding damage by AcST-3-infected *T. ni* larvae.

METHODS. Third instar larvae were fed individually overnight with a plug of artificial diet inoculated with 10^4 polyhedra. This dose ensured 100% mortality. After feeding, the larvae were enclosed in groups of five with cabbage leaves from plants that had been grown in a controlled environment. Before feeding, the leaves had been soaked overnight to ensure water saturation. After two days of feeding new foliage was introduced. All larvae were dead after a further 3 days when the photograph was taken. The experiment was conducted at 23 °C. Control uninfected larvae fed on cabbage leaves pupated normally. In other experiments, the surface area of the leaves was measured before, and after feeding to the larvae using a Portable Area Meter: Model L1 3000 (Lambda Instruments).

We have modified a baculovirus insecticide by inserting a gene encoding an insect-specific toxin so that it reduces survival time of infected insects and the amount of host plant damage. The virus can still produce polyhedra, making a realistic proposition for field use, where it should cause less crop damage than normal, unmodified baculovirus insecticides. □

Received 3 April; accepted 17 May 1991.

1. Cunningham, J. C. in *Microbial and Viral Insecticides* (ed. Kurstak, E.) 335-386 (Dekker, New York, 1982).
2. Entwistle, P. F. & Evans, H. F. in *Comprehensive Insect Physiology Biochemistry and Pharmacology* (eds Kerkut, G. A. & Gilbert, L. I.) 347-412 (Pergamon, Oxford, 1985).
3. Huber, J. in *The Biology of Baculoviruses II. Practical Applications for Insect Control* (eds Granados, R. R. & Federici, B. A.) 181-202 (CRC, Boca Raton, 1986).
4. Evans, H. F. & Entwistle, P. F. in *Epizootiology of Insect Diseases* (eds Fuxa, J. R. & Tanada, Y.) 257-322 (Wiley, New York, 1987).
5. Smith, G. E., Summers, M. D. & Fraser, M. J. *Molec. cell. Biol.* **3**, 2156-2165 (1983).
6. Pennock, G. D., Shoemaker, C. & Miller, L. K. *Molec. cell. Biol.* **4**, 399-406 (1984).
7. Carbonell, L. F., Hodge, M. R., Tomalski, M. D. & Miller, L. K. *Gene* **73**, 409-418 (1988).
8. Maeda, S. *Biochem. biophys. Res. Commun.* **165**, 1177-1183 (1989).
9. Merryweather, A. T. et al. *J. gen. Virol.* **71**, 1535-1544 (1990).
10. Martens, J. W. M. et al. *Appl. envir. Microbiol.* **56**, 2748-2754 (1990).
11. Hammock, B. D., Bonning, B. C., Possee, R. D., Hanzlik, T. N. & Maeda, S. *Nature* **344**, 458-461 (1990).
12. Zlotkin, E., Rochat, H., Kopeyan, C., Miranda, F. & Lissitzky, S. *Biochimie* **53**, 1073-1078 (1971).
13. Walther, C., Zlotkin, E. & Rathmeyer, W. *J. Insect Physiol.* **22**, 1187-1194 (1976).
14. Teitelbaum, Z., Lazarovici, P. & Zlotkin, E. *Insect Biochem.* **9**, 343-346 (1979).
15. Darbon, H., Zlotkin, E., Kopeyan, E., Van Rietschoten, J. & Rochat, H. *Int. J. Peptide Protein Res.* **20**, 320-330 (1982).
16. Bougis, P. E., Rochat, H. & Smith, L. A. *J. biol. Chem.* **264**, 19259-19265 (1989).
17. Zlotkin, E. *Insect Biochem.* **13**, 219-236 (1983).
18. de Dianous, S., Hourau, F. & Rochat, H. *Toxicol.* **25**, 411-417 (1987).
19. Weyer, U., Knight, S. & Possee, R. D. *J. gen. Virol.* **71**, 1525-1534 (1990).
20. Whitford, M., Stewart, S., Kuzio, J. & Faulkner, P. J. *Virol.* **63**, 1393-1399 (1989).
21. Matsuura, Y., Possee, R. D., Overton, H. A. & Bishop, D. H. L. (1987). *J. gen. Virol.* **68**, 1233-1250.
22. Kunkel, T. A. *Proc. natn. Acad. Sci. U.S.A.* **83**, 488-492 (1985).
23. Livingstone, C. & Jones, I. *Nucleic Acid Res.* **17**, 2366 (1989).
24. Kitts, P. A., Ayres, D. & Possee, R. D. *Nucleic Acids Res.* **18**, 5667-5672 (1990).
25. Possee, R. D. *Virus Res.* **5**, 43-59 (1986).
26. Cook, R. F., Avery, R. J. & Dimmock, N. J. *Infect. Immunity* **25**, 396-402 (1977).
27. Finney, D. J. *Probit Analysis* 3rd edn (Cambridge University Press, UK, 1971).
28. Hughes, P. R. & Wood, H. A. *J. Invert. Pathol.* **37**, 154-159 (1981).
29. Hughes, P. R., van Beek, N. A. M. & Wood, H. A. *J. Invert. Pathol.* **48**, 187-192 (1986).
30. Harap, K. A., Payne, C. C. & Robertson, J. S. *Virology* **79**, 14-31 (1977).

CORRECTION

An explanation for the protective effect of the MHC class II I-E molecule in murine diabetes

Eva-Pia Reich, Robert S. Sherwin, Osami Kanagawa & Charles A. Janeway Jr

Nature **342**, 326-328 (1989)

A MAJOR finding in the above report in the 28 September 1989 issue was that cloned T-cell lines that could transfer insulinitis to young irradiated recipient mice bore antigen receptors encoded by *V β 5* genes. We now have convincing data that the receptors are not encoded by *V β 5* genes. The original finding was based on samples of a monoclonal antibody specific for *V β 5*. The specificity of this antibody has been validated in numerous studies. But several batches of this antibody produced after the paper was published failed to stain our cloned T-cells. We believe that the batch used in the experiments we reported contained a contaminant that led to spurious results. Recent molecular analysis shows that the β chains of the T-cell receptors in the clones of interest are encoded by *V β 1* genes. A detailed characterization of these receptors will be reported when the molecular analysis is completed. The other findings in the paper have been validated in multiple experiments, and we have no reason to doubt them. We deeply regret any negative effects this error may have had on research in this important area. □

nature
is available in
microform.



University Microfilms

International reproduces this publication in microform: microfiche and 16mm or 35mm film. For information about this publication or any of the more than 13,000 titles we offer, complete and mail the coupon to: University Microfilms International, 300 N. Zeeb Road, Ann Arbor, MI 48106. Call us toll-free for an immediate response: 800-521-3044. Or call collect in Michigan, Alaska and Hawaii: 313-761-4700.

☐ Please send information about these titles:

Name _____

Company/Institution _____

Address _____

City _____

State _____ Zip _____

Phone (____) _____

University
Microfilms
International

Human genome analysis system

Isao Endo, Eiichi Soeda, Yasubumi Murakami and Katsuo Nishi

An automated DNA sequencing system, the collaborative effort of Japanese scientists and engineers, which has a potential output of up to 108,000 bases per day is described.

SINCE 1988, the human genome analysis project, entitled Studies on Gene Structure, has been promoted by the Science and Technology Agency (STA) of Japan. By developing this project, Japanese scientists and engineers hope to play their part in not only the elucidation of human biology and medicine but also in the development of basic biology.

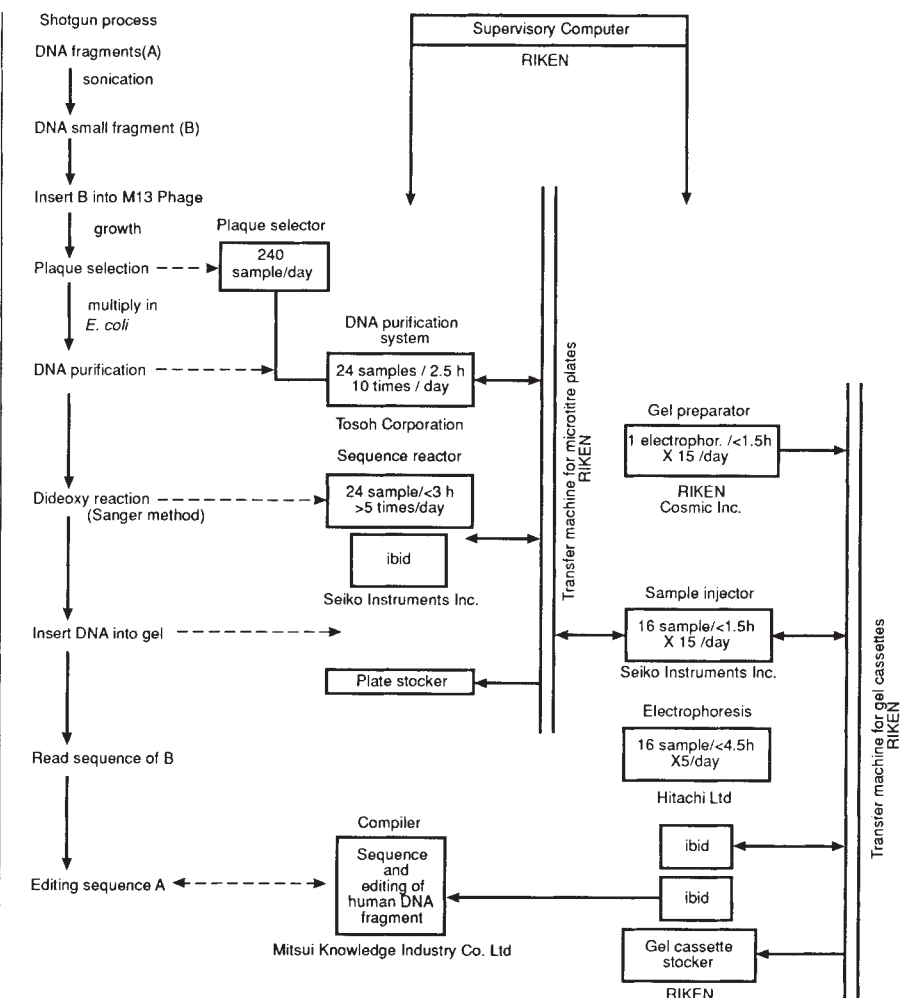
The aims of this project are twofold. First, to develop an automated rapid analysis system for sequencing the human genome and, second, to elucidate preparation techniques for the genome sequences, which include the preparation of genome libraries, mapping techniques and data base systems.

Scientists and engineers from 12 research facilities, which include universities, government institutions and private companies, have participated in this project and about ¥600 million (\$4.5 million) has already been founded by STA. Research work targeted at accomplishing the first objective of the project has been led by the authors. The second part, by Professor Yohji Ikawa. The article describes an automated human genome analysis system (HUGA-I), which is expected to play a significant part in Japan's efforts to sequence the human genome.

The process of determining the nucleotide sequence of large DNA molecules, like the human genes, is tedious and time-consuming, and yet requires our unstinting energy. It would be a significant contribution to the development of medical science and biology if these problems could be solved by automating, accelerating and taking a systematic approach to the conventional process, an idea that was first proposed by Wada¹ in 1987. He and his team of scientists and engineers have developed several automated machines with the backing of the Special Coordination Fund for the Promotion of Science and Technology (1981–1984) and the Cancer Research Fund (1985–1987) under the auspices of STA. These funds allowed for the completion of machines like the DNA extractor (Seiko Instruments, Inc., Matsudo-Shi, Chiba), dideoxy reactor (ibid), gel-forming machine (Fuji Film Co. Ltd, Asaka-Shi, Saitama) and the fluorescent sequencer (Hitachi Ltd, Chiyoda-Ku, Tokyo).

System set-up

Our prototype model is designed to accelerate the functions of the machines already developed and to combine the whole sequencing process according to the modified shotgun method that was proposed by Deininger². It consists of the following main unit processes: construction of human DNA



Japan's human genome analysis system (HUGA-I).

libraries, cloning of DNA fragments, purification of these fragments, dideoxy reactions, injection of DNA molecules from the dideoxy reactor into the gel, sequencing gel electrophoresis and editing numbers of short sequences into the original DNA sequence.

Processes such as the construction of DNA libraries, sonication of the test DNA (~ 100 kilobases) into smaller sizes (~ 1 Kb), cloning and amplification of the smaller DNA fragments using M13 phages and *Escherichia coli* could not be automated because these processes are unpredictable. However, plaque selection from the plate and the transfer of the microorganisms into test tubes are routine operations, and Seiko Instruments has automated this operation by use of a robotic arm. The downstream processes, from the purification of DNA fragments to editing the DNA sequences, have been automated by the use of two kinds of transfer robots: one for the transfer of micro-

titre plates and the other for the transfer of the gel cassette. A schematic diagram of the system is shown in the figure.

System throughput

The concept of the system is as follows: if it could be used to elucidate 100 kb of human DNA sequence per day and could operate continuously for 300 days a year, then it would be possible to read the three billion base pairs of human DNA sequence within 10 years by using 10 lines of this system. A fluorescent sequencer that can read 450 bases per sample within 4.5 hours has already been developed by Hitachi. The capacity of the slab gel is 16 samples and so if three sequencers are operated simultaneously, each for five cycles a day, then 108,000 bases could be analysed in a day.

450 bases per sample × 16 samples per sequencer × 3 sequencers per cycle × 5 cycles per day = 108,000 bases per day

In order to operate Hitachi's sequencer efficiently, operating schedules and capacities of upstream processes were determined according to the flow diagram in the figure. The Institute of Physical and Chemical Research (RIKEN), in collaboration with Cosmic, Inc. (Minato-Ku, Tokyo), has developed a gel preparation machine (Gel Preparator) in which the slab gel (polyacrylamide gel, 0.3- mm thickness) is solidified and cleaned. A sample injection robot that can automatically inject 16 samples into the top of a slab gel cassette in 1.5 hours has been developed by Seiko Instruments. After all the samples are injected into the slab gel, electrophoresis takes place for one hour. The cassette is then transferred to the sequencer by the transfer robot developed at RIKEN.

Meanwhile, samples are prepared as follows: human DNA cloned into a cosmid vector is broken into smaller fragments by sonication, and then cloned using M13 phage and *E. coli* systems. Plaques that contain human DNA fragments are automatically selected and inoculated into test tubes by a plaque selector (Seiko Instruments). This machine can provide 240 samples per day. The DNA fragments are then amplified in test tubes for six hours and purified using a DNA purification system that was developed by Tosoh Corporation (Minato-Ku, Tokyo). With this membrane-based method of purification, DNA fragments, *E. coli* cells and phages are successively filtered, first by a microfiltration membrane and then by an ultrafiltration membrane. The DNA fragments obtained are redissolved in a Tris-EDTA solution and then one microlitre of this solution is pipetted automatically into 24 designated wells of a 96-well microtitre plate. The procedure takes 2.5 hours to complete, which means that the machine can be set to repeat the process at least nine times per day.

After completion of the pipetting routine, another transfer machine, also developed at RIKEN, picks up the microtitre plate and transfers it to the dideoxy reactor (Seiko Instruments)³. This reactor, the so called 'chemical robot', can perform a complete set of Sanger's sequencing reactions. The main functions of this machine are quantitative addition and dispensing of reagents, mixing, heating and cooling — all within the unit. Using this machine, the dideoxy sequencing reactions for 24 samples take three hours. Consequently, the machine can be set to repeat the process more than five times per day. In order to provide DNA molecules to the sequencers, two of these machines have been installed in the system. After the Sanger reactions are complete, the transfer machine again picks up the microtitre plate and hands it on to the sample injector, which was mentioned earlier.

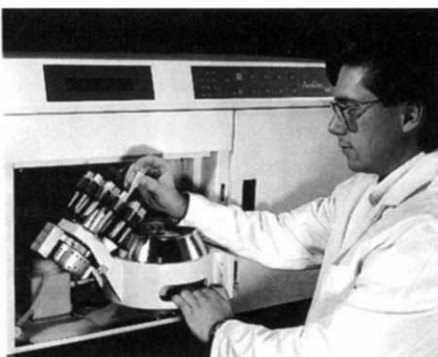
On the software side

RIKEN, in collaboration with Mitsui Knowledge Industry Co. Ltd (Chiyoda-Ku, Tokyo), has developed a computer program

DNA details

Handy helpers in this week's issue include a dual-purpose device for vertical electrophoresis and electroblotting, a triple-tray gel box and news of a new patent sequence data bank.

AN additional seven **DNA isolation protocols** are available for the AutoGen 540, an instrument that has broad applicability in genetic engineering, DNA sequencing and mapping, and DNA probe diagnostics, which is manufactured by a company of the same name (*Reader Service No. 101*).



AutoGen 540 for automated DNA isolation.

Autogen claims that the 540 is the only fully automated instrument chemistry system with the flexibility to isolate plasmids, cosmids, M13, lambda and genomic DNA from a wide variety of biological sources. The system is designed to process up to 12 samples simultaneously in cycles times that range from 60 to 130 minutes. In addition to user programmability, the most commonly requested DNA isolation procedures are available on floppy disk. As protocol development is an ongoing process at AutoGen, the company supplies users with regular

updates. Recently introduced protocols include: a genomic extraction procedure for the isolation of genomic DNA from blood, lymphocyte preparations, mouse tails and other tissue suspensions; an improved plasmid/cosmid protocol that produces DNA suitable for *in situ* hybridization; a protocol for M13 and lambda single-stranded DNA; and a DNA isolation protocol that produces DNA suitable for multiplex sequencing, fluorescent sequencing and mapping.

Gel box assortment

Lifecodes, a supplier of DNA probes, accessories and supplies, has introduced a new **triple-tier gel box**, which is designed to save processing time and space within the laboratory (*Reader Service No. 102*). Each gel box includes three heavy-duty, UV-transparent gel trays, electrodes with connector leads and buffer recirculating ports. The triple-tray gel box is available in two sizes: a long-gel set (12 × 28 cm) and a medium-gel set (12 × 21.5 cm). Each costs \$600 (US).

Both gel fractionation and electrotransfer can be performed using only one device with the **dual-purpose mini-vertical electrophoresis system** from Schleicher and Schuell (*Reader Service No. 103*). The conversion from gel to blot is made in one step by removing the mini-electrophoresis assembly and inserting the miniblitter assembly, which incorporates ceramic electrodes for more even transfer and cleaner electroblots, S & S

for sequence assembly and editing. In the assembly of random sequence fragments, the computer links up matching portions of the fragments to reconstruct the original complete sequence. Where ambiguous data exist, the computer leaves gaps in the sequence. To increase the accuracy of this approach, fragments that contain the same portions of the original sequence are automatically read from different directions and different locations approximately four times. It should be noted that inevitable errors occurring in the raw sequence data will be corrected or reduced by overlapping these sequences. By combining our experiences, an improved program for compilation of the sequence is being prepared, which will be applicable to any VAX computer (DEC Ltd, Maynard, Massachusetts).

In this fiscal year (1991 to 1992), the performance of HUGA-I will be evaluated using DNA cosmids. Even though it is the intention to use HUGA-I to elucidate the human genome, starting in April 1992, its potential

is by no means limited to this alone. HUGA-I could be used to map out the genetic sequence of any living organism. It must also be pointed out that in order for additional progress to be made, development of the hardware must go hand-in-hand with the methodology. In this way, scientists, freed from such time-consuming monotonous work, will be able to concentrate on conducting original research. □

Isao Endo, Eiichi Soeda, Yasubumi Murakami and Katsuo Nishi are in the Division of Genome Research, the Institute of Physical and Chemical Research (RIKEN), Koyadai, Tsukuba-shi, Ibaraki, 305, Japan. The authors thank the STA, the contributing companies and Professor Akiyoshi Wada for their contributions to this project. For more information, fill in reader service number 100.

1. Wada, A. *Nature* **305**, 193 (1984).
2. Deininger, P.L. *Analyt. Biochem.* **129**, 216-218 (1983).
3. Sakabe, M. et al. *Rev. Sci. Instrum.* **61** 1966-1970 (1990).



One tank, two techniques — see S & S.

says. When using the S & S set-up, either technique requires less than one hour running time for complete separation or blotting. In addition, two gels can be run simultaneously. The \$695 (US) system is designed for use with S & S's precast polyacrylamide gels. Gels are available for both protein and nucleic acid (Tris-borate-EDTA) electrophoresis in a wide variety of isocratic and gradient formats.

Statagene's new ST/RIDE device is designed for the **resolution of megabase-size DNA** by simultaneous tangential, rectangular decussate electrophoresis (*Reader Service No. 104*). With the \$8,995 (US)



ST/RIDE system separates megabase DNA.

ST/RIDE device, the gel stands vertically in an acrylic tank with six replaceable electrodes positioned around the gel in a hexagonal pattern that is perpendicular to the plane of the gel. The unique electrode geometry generates a homogenous electrical field over the entire area of the gel creating straight-lane separations of DNA spanning a wide range of molecular sizes, Statagene says. The method used is based upon the application of a constant electrical field from top to bottom through a vertical gel in combination with a perpendicular field that alternates

polarity in a cyclic fashion. Variation of pulse time, field strength and pulse vector angle can be achieved using a microprocessor-controlled power supply. The ST/RIDE device has been used for the separation of *Saccharomyces cerevisiae* chromosomes, lambda concatamers and *Hind* III-digested lambda DNA.

For the electrophoretic separation of **high molecular weight DNA fragments or entire chromosomes**, Biometra recommends the Rotaphor 22 system with its crescent-shaped rotating electrodes (*Reader Service No. 105*). Separations of lambda phage can be achieved in 19 hours using the Rotaphor 22 device and the rotating field electrophoresis technique; separations of *Saccharomyces cerevisiae* and *S. pombe* take 24 and 48 hours, respectively. The chamber provides 20 x 20 cm of gel area. By turning the gel tray through 90°, two-dimensional electrophoresis can be performed. The controller is equipped with a memory to store up to 10 run programmes; nine of these in a non-volatile memory. The Rotaphor R22 system, which includes the chamber, safety lid, rotating electrodes, gel tray, self-sealing gel frame, comb (16 wells), sample mould and the latest R22 controller, costs £4,150 (UK).

DNA amplification devices

Two new **thermal cycling units**, the Biosycler Block and the Biosycler Dual Block, are hot off the Bios production line (*Reader Service No. 106*). The units vary in terms of capacity: the Biosycler Block can hold 29 microfuge tubes, whereas the Biosycler Dual Block holds double that quantity (58 microfuge tubes, 29 per block). Both units feature the same temperature control mechanism found in the company's Biosycler Oven, which is provided by a thermal probe that mimics actual sample temperature and controls the transition between time/temperature steps. A maximum of 20 programs can be linked. The Biosycler Block can be programmed to run a different program in each block, providing two instruments in one. Alternatively, both blocks can be set to run the same programme simultaneously. The \$3,295 (US) Biosycler Block includes parallel printer and chart recorder interfaces. However, the \$5,435 (US) Biosycler Dual Block is supplied with a built-in chart



Biosycler Block thermal cycling unit.

ADVERTISEMENTS

DNA

Around the World

\$4.50 a base
(no set-up)

Includes:

- 500-1500 ugs
- 15-45 O.D.S.
- 24-72 hour delivery
- Amplification ready
- Gel picture
- Quality guaranteed products

We are a full service and design DNA lab.

1-800-227-0627
Fax 1-214-420-0442

BIO-SYNTHESIS

Reader Service No.18

SINGER

INSTRUMENTS

The original

SINGER MK I

MICROMANIPULATOR

single or in pairs

FOR

- XENOPUS OOCYTE INJECTION
- PIPETTING OF CELLS IN SUSPENSION
- MICRODISSECTION AND MICRO SURGERY

INT. SALES: +44.984.40226
FAX: +44.984.41166 for details

Reader Service No. 50

Synthetic Oligonucleotides

Custom made probes and primers,
synthesised, purified and shipped in 3 days

£3.50 per base

(All inclusive)
0.2 µM scale (approx. 500 µg)

Modified bases, biotin or digoxigenin labelling,
data base analysis and full design facilities

Molecular Medicine Unit,
King's College School of Medicine & Dentistry,
Denmark Hill, London SE5 8RX

Tel. XX-44-(0)71 - 326 3126
FAX XX-44-(0)71 - 733 3877

Reader Service No.56

recorder. Graphical output can be obtained by connecting the unit to a printer.

The BioPhase **programmable thermal cycler** from BioExcellence provides users with rapid temperature transition under software control from below ambient to above 100°C (*Reader Service No. 107*). The solid-state technology used in the cycler incorporates a unique heating system, with well insulated and protected Peltier cells for rapid and efficient cooling, BioExcellence says. Menu-driven software allows the user to create up to 50 user-defined run programs. Steps can be set in one-second intervals to virtually any duration. Continuous monitoring under microprocessor control is designed

ADVERTISEMENTS

Reader Service No.79

TIB MOLBIOL BERLIN
CUSTOM DNA SYNTHESIS

5,- DM per base

DM 95 - per oligo

FAX (0049) 30 6261960

TRITIATED & ULTRA PURE PEPTIDES

Reader Service No.15

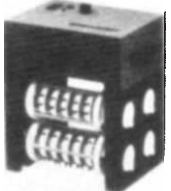
An increasingly recognized problem is the purity of peptides used in biological studies. We offer synthesis of peptides including chemical synthesis of tritiated peptides of a very high purity.

SERTEK

Haraldsgade 68, DK-2100 Denmark
Tel: +45-3120 2554; Fax +45-3181 7833

Reader Service No.49

BIANORM
EQUILIBRIUM DIALYSIS SYSTEM



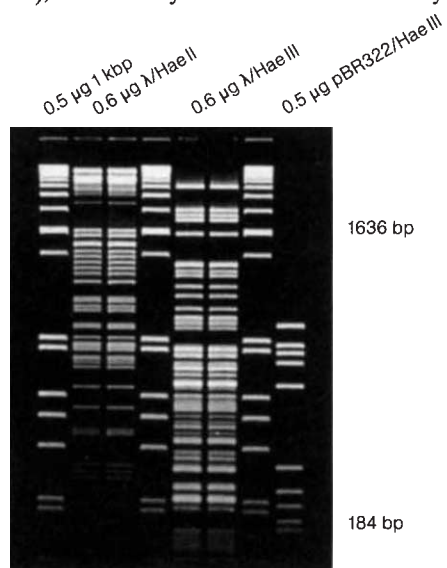
The gentlest method to quantify the binding parameters and thermodynamic data between ligands and biopolymers. Results come closest to the actual values. Sample volume from 0.2 ml to 5.0 ml. Up to 20 experiments in parallel.

BIANORM
Geräte
P.O. Box 550726
D-4000 Mönchen 65,
FRG

to ensure rapid temperature equilibrium at each preset step. For a permanent record of the run conditions, the instrument can be hooked up to a printer.

Gels, gels, gels

Elchrom of Switzerland is offering precast poly(NAT) gels, which are designed for the analysis of double-stranded DNA by submarine electrophoresis (*Reader Service No. 108*). The three **ready-to-use synthetic gels**, which are available in gel strengths of 6%, 9% and 12%, provide high resolution in the range of 30 to 6,000 base pairs, even in samples with more than 100 fragments (λ /Hae III), Elchrom says. The detection sensitivity



Precast 9% poly(NAT) gel.

is less than 0.5 ng of a 1-kbp fragment. In addition, the plot of log DNA size versus migration distance is linear in the optimal separation range of each gel, says the company. Poly(NAT)gels, which measure 62 x 92 mm, are sold five gels to the box and in unit quantities of one and ten boxes for \$60 (US) and \$510 (US), respectively.

BDH has upgraded its entire range of **acrylamides** for applications in biochemistry and molecular biology (*Reader Service No. 109*). Acrylamide grade 1 has been upgraded by the addition of new test parameters and by the tightening of current specifications to create the new Grade 1-Molecular Biology Grade. This grade of acrylamide is intended for molecular biologists requiring stringent standards of purity. Acrylamide Electran, which is designed for use in most analytical procedures, is suitable for all forms of protein analysis and many aspects of nucleic acid research. A new grade of acrylamide, Acrylamide Electran Routine, is intended for routine screening applications, such as analysing column fractions and checking process steps.

Sequence analysis

The types of sequences overlooked when searching protein and nucleic acid data

banks are often those held by patent offices — most of which never appear in the public data banks such as GenBank, EMBL, SWISS-PROT and PIR, or in the scientific literature. IntelliGenetics, in collaboration with Derwent, has collected this information together into GENESEQ, a new **data bank of nucleic acid and protein sequences found in patent publications** (*Reader Service No. 110*). The company expected that by June of this year, the number of sequences unique to GENESEQ would top 10,000. The GENESEQ data bank is available for Sun workstations and VAX computers, and via the IntelliGenetics Timesharing System. Accounts with this on-line system, which provides users with access to the IntelliGenetics Suite, GENESEQ, GenBank, EMBL, SWISSPROT and PIR, can be set up for \$200 (US).

International Biotechnologies has released Version 3.5 of its **MacVector sequence analysis software program** (*Reader Service No. 111*). So what's new you ask. Listening to comments of current MacVector users, IBI says it has increased the speed of several analyses, some by as much as five-fold. And, users now have the option of using any one of ten different genetic codes when translating, reverse translating, or when comparing DNA and protein sequences to one another. A separate menu option has been added for comparing multiple sequences to either a DNA or protein query. Users of MacVector 3.5 can now access the GenBank nucleic acid data base from a CD-ROM disk, which is updated four times a year. In addition, MacVector 3.5 features



IBI's sequence analysis software upgrade.

boolean data base searches, negative sequence numbering, more comprehensive graphics, extended line lengths from 60 characters per line to 225 characters, new sequence editors and an intuitive restriction enzyme editor. Existing MacVector users (Version 3.0) can obtain this upgrade for \$125 (US). Version 3.5 for new users costs \$2,495 (US). □

These notes are compiled by Diane Gershon from information provided by the manufacturers. To obtain further details, use the reader service card bound inside the journal. Prices quoted are sometimes nominal, and apply only within the country indicated.